

Copules dynamiques pour séries chronologiques

par

Antoine Bergeron

Mémoire présenté au Département de mathématiques
en vue de l'obtention du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, juin 2020

Le 4 juin

*le jury a accepté le mémoire de Monsieur Antoine Bergeron
dans sa version finale.*

Membres du jury

Professeur Taoufik Bouezmarni
Directeur de recherche
Département de Mathématiques

Professeur Carole Beaulieu
Codirectrice de recherche
Département de Biologie

Professeur Pierre Dutilleul
Codirecteur de recherche
McGill University, Department of Mathematics and Statistics

Professeur Patrick Richard
Membre externe
Département d'économie

Professeur Éric Marchand
Président-rapporteur
Département de Mathématiques

À ma famille,
qui m'a soutenu jusqu'à la fin,
ainsi qu'à ma conjointe
Émilie.

SOMMAIRE

La structure de dépendance que possède une série chronologique est un sujet d'étude populaire depuis très longtemps. Cependant, plusieurs approches se limitent à considérer une dépendance fixe. Ce mémoire se veut une ouverture à l'application d'une structure de dépendance « variant » dans le temps pour une série chronologique univariée définie par un processus markovien d'ordre 1 à l'aide de copule.

Le chapitre 1 concerne les principes de base de l'analyse chronologique. De nombreux modèles applicables aux séries temporelles ainsi que leurs propriétés sont présentés.

Le chapitre 2 introduit les copules, un outil maintenant très utilisé pour l'étude de la structure de dépendance. On y trouve une introduction à la fonction copule, les principales classes et familles ainsi que différentes méthodes d'estimation pour celles-ci.

Le chapitre 3 présente l'application des copules dans les séries chronologiques. La dépendance de plusieurs modèles paramétrique des séries chronologiques est d'abord étudiée. On scrute ensuite la dépendance dans une seule série temporelle, lorsque celle-ci est fixe ou varie dans le temps. Des algorithmes de simulation sont aussi présentés.

Le chapitre 4 rapporte des résultats de simulations distribués entre des applications théoriques, afin d'évaluer le comportement des estimateurs. Une application empirique utilisant les séries du prix du blé de cinq pays est aussi présentée.

REMERCIEMENTS

Je tiens d’abord à remercier mon directeur de maîtrise, M. Taoufik Bouezmarni, pour ces nombreuses explications et son expertise dans le domaine des copules. Je tiens aussi à remercier mon directeur de recherche adjoint, M. Pierre Dutilleul, pour ses conseils inestimables en statistiques ainsi que son aide afin de mieux structurer mes idées. Ce fut un plaisir de discuter et d’échanger avec vous deux pendant ces trois ans que constitue ma maîtrise. Je veux aussi souligner l’apport de Carole Beaulieu pour les applications liées à la biologie.

Je veux également remercier mes collègues du Département de mathématiques. Un merci spécial à Olivier Germain, Dominic Desjardin Côté et Denis Langford.

Finalement, un merci tout spécial à ma famille, ma copine ainsi qu’à mes amis pour leur soutien tout au long de ma maîtrise.

Antoine Bergeron
Sherbrooke, juin 2020

TABLE DES MATIÈRES

SOMMAIRE	iii
REMERCIEMENTS	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	viii
LISTE DES FIGURES	xi
INTRODUCTION	1
CHAPITRE 1 — Les séries chronologiques	5
1.1 Les types de modèles	11
1.2 Processus à moyenne mobile (MA)	12
1.3 Processus autorégressif (AR)	14
1.4 Processus autorégressif et moyenne mobile (ARMA)	17

1.5	Processus autorégressif conditionnellement hétéroscédastique (généralisé)	19
1.6	Processus autorégressif et moyenne mobile intégré (ARIMA)	23
1.7	Processus autorégressif et moyenne mobile intégré fractionnaire (ARFIMA)	25
CHAPITRE 2 — Les copules		26
2.1	Définitions et propriétés	26
2.2	Mesures de dépendance	28
2.3	Classes et familles de copules	31
2.3.1	Classe elliptique	31
2.3.2	Classe archimédienne	37
2.4	Estimation des copules	41
CHAPITRE 3 — Copules statique et dynamique		49
3.1	La copule multivariée	49
3.1.1	Algorithmes de simulation	55
3.2	Copule bivariable à paramètre constant pour série chronologique univariée	58
3.3	Copule bivariable à paramètre dynamique pour série chronologique univariée	63
3.3.1	Algorithme de simulation	72
CHAPITRE 4 — Applications		75
4.1	Applications théoriques	76
4.2	Application empirique	87

CONCLUSION	94
BIBLIOGRAPHIE	96

LISTE DES TABLEAUX

2.1	Tau de Kendall et tau de Kendall inverse pour chacune des copules d'intérêt, avec $D_k(x) = \frac{k}{x} \int_0^x \frac{s^k}{e^s - 1} ds$	47
3.1	Tau de Kendall et inverse du tau de Kendall dynamique pour chacune des copules d'intérêt, avec $D_k(x) = \frac{k}{x} \int_0^x \frac{s^k}{e^s - 1} ds$	65
4.1	Intervalles de confiance (I.C.) pour les paramètres β_0 et β_1 dans le cas de la copule gaussienne avec $\beta_0 = -1$ fixe et $N = 500$ séries de $T = 500$ observations	77
4.2	Intervalles de confiance (I.C.) pour les paramètres β_0 et β_1 dans le cas de la copule gaussienne avec $\beta_0 = -1$ fixe et $N = 500$ séries de $T = 200$ observations	78
4.3	Intervalles de confiance (I.C.) pour les paramètres β_0 et β_1 dans le cas de la copule gaussienne avec $\beta_0 = -1$ fixe et $N = 500$ séries de $T = 100$ observations	79
4.4	Estimation des paramètres pour différentes familles de copule obtenue à l'aide de $N = 500$ séries de $T = 500$ observations, pour $\beta_0 = -0.5$ et $\beta_1 = 0.2$	80

4.5	Estimation des paramètres pour différentes familles de copule obtenue à l'aide de $N = 500$ séries de $T = 500$ observations, pour $\beta_0 = -0.3$ et $\beta_1 = -0.2$	81
4.6	Estimation pour différentes familles de copule obtenue à l'aide de $N = 500$ séries de $T = 150$ observations, pour des paramètres $\beta_0 = -0.5$ et $\beta_1 = 0.2$	83
4.7	Estimation pour différentes familles de copule obtenue à l'aide de $N = 500$ séries de $T = 150$ observations, pour des paramètres $\beta_0 = -0.3$ et $\beta_1 = -0.2$	83
4.8	Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé de l'Australie, $T = 104$	88
4.9	Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé du Royaume Uni, $T = 109$	88
4.10	Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé du Kazakhstan, $T = 130$	89
4.11	Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé du Canada, $T = 141$	89
4.12	Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé de la région de la Mer Noire, $T = 76$	89
4.13	Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé de l'Australie, $T = 104$	90
4.14	Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé du Royaume-Uni, $T = 109$	91
4.15	Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé du Kazakhstan, $T = 130$	92

4.16 Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé du Canada, $T = 141$	92
4.17 Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé de la Mer Noire, $T = 76$	92

LISTE DES FIGURES

1.1	Graphique d'un processus MA(4) de 500 observations avec paramètres $\theta_1 = 0.8$, $\theta_2 = -0.5$, $\theta_3 = 0.3$, $\theta_4 = 0.2$, $\sigma_\epsilon^2 = 1$, $\mu = 0$ et de sa fonction d'autocorrélation	14
1.2	Graphique d'un processus AR(3) de 500 observations avec paramètres $\phi_1 = 0.8$, $\phi_2 = 0.2$, $\phi_3 = -0.5$, $\sigma_\epsilon^2 = 1$, $\nu = 0$ et de sa fonction d'autocorrélation partielle	17
1.3	Graphique d'un processus ARMA(3,1) de 500 observations avec paramètres $\phi_1 = 0.8$, $\phi_2 = -0.2$, $\phi_3 = 0.3$, $\theta_1 = 0.8$, $\sigma_\epsilon^2 = 1$ et $\mu = 0$	20
1.4	Graphiques des fonctions d'autocorrélation et d'autocorrélation partielle associées au processus ARMA(3,1) de la figure 1.3	20
1.5	Graphique d'un processus ARCH(2) de 500 observations avec paramètres $\alpha_0 = 0$, $\alpha_1 = 0.2$, $\alpha_2 = 0.4$	22
1.6	Graphique d'un processus GARCH(1,1) de 500 observations avec paramètres $\alpha_0 = 0$, $\alpha_1 = 0.1$, $\beta_1 = 0.8$	23
1.7	Graphiques du nombre de passagers par vol international à gauche et de sa différentiation d'ordre 1 à droite	24

2.1	Graphiques de la densité et nuages de points ($n = 1000$) pour des copules gaussiennes bivariées avec paramètre $\rho = -0.8, 0$ et 0.8	33
2.2	Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Student bivariées avec paramètre $\rho = -0.8, 0, 0.8$ et 3 degrés de liberté (ddl).	35
2.3	Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Gumbel bivariées avec paramètre $\theta = 4$ et 8	38
2.4	Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Clayton bivariées avec paramètre $\theta = 4$ et 8	39
2.5	Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Frank bivariées avec paramètre $\theta = 4$ et 8	40
4.1	Graphiques de la valeur du tau de Kendall pour $\tau_t = \exp\left(\beta_0 + \beta_1 \cdot \frac{t}{T}\right)$. À gauche, $\beta_0 = -0.5, \beta_1 = 0.2$ et $T = 500$. À droite, $\beta_0 = -0.3, \beta_1 = -0.2$ et $T = 500$	82
4.2	Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule gaussienne du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$	84
4.3	Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Student du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$	85
4.4	Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Gumbel du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$	85

4.5	Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Clayton du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$	86
4.6	Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Frank du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$	86
4.7	Valeurs du tau de Kendall pour $\tau_t = \exp\left(\beta_0 + \beta_1 \cdot \frac{t}{T}\right)$, avec β_0 et β_1 les valeurs estimées pour (A), la copule de Student (Kazakhstan) et (B), la copule de Frank (Royaume-Uni).	91

INTRODUCTION

En statistique, l'étude d'évènements est chose très courante. Ces évènements, souvent représentés par un ensemble de variables, peuvent être étudiés de plusieurs façons. L'une d'elles est via l'analyse de séries temporelles/chronologiques/chroniques. Une série temporelle est composée par définition d'observations faites pour une variable statistique variant dans le temps. Outre le fait de pouvoir les décrire et celui de pouvoir expliquer leur variation à l'aide d'un modèle, l'utilité des séries temporelles est de pouvoir prédire des observations futures. La valeur d'une action en finance et la concentration de gaz carbonique dans l'atmosphère sont des exemples bien connus de séries chronologiques. Le 20^e siècle fut une période fertile pour la proposition de modèles en analyse chronologique. Un modèle de base est de définir la variable au temps t en fonction des variables aux temps $(t - 1), (t - 2), \dots, (t - p)$, en plus d'un bruit blanc ou d'une constante. Ce modèle fut proposé par Yule [Yul27] et Walker [Wal31], et porte le nom de modèle autorégressif (AR) d'ordre p . Quelques décennies plus tard, Sims [Sim80] proposa le modèle vectoriel autorégressif (VAR) pour les cas multivariés, où plusieurs séries chronologiques sont étudiées pour leur interdépendance. Whittle [Whi51] propose ensuite le modèle de la moyenne mobile (MA) dans lequel une variable au temps t est expliquée par d'autres variables à des temps antérieurs. Spécifiquement, une variable est expliquée par un modèle additif constitué d'une moyenne (celle de la série) et de bruits blancs (ou innovations) aux temps présent et

précédents, c'est-à-dire, $t, (t-1), (t-2), \dots, (t-q)$. Dans ce contexte, q est appelé l'ordre de la moyenne mobile. L'utilisation conjointe des modèles AR et MA mène au processus ARMA(p, q), portant le nom de *Autoregressive-moving-average model*. L'ordre p est lié au modèle AR alors que l'ordre q est lié au modèle MA. Dans ce modèle, la variable d'intérêt au temps t est expliquée par la même variable aux temps

$(t-1), (t-2), \dots, (t-p)$, ainsi que par des bruits blancs aux temps $t, (t-1), (t-2), \dots, (t-q)$. De plus, mis en valeur au début des années 1970 par Box et Jenkins, le modèle ARIMA(p, d, q) *Autoregressive Integrated Moving Average* permet de prendre en compte une tendance d'ordre d . Dans ce modèle, le p représente l'ordre du processus AR, le q représente l'ordre du processus MA, et le d représente l'ordre de la différenciation. La différenciation signifie que la série originale est remplacée par la série des différences adjacentes. Ces trois modèles sont basés sur la linéarité. Par contre,

il se peut qu'une série chronologique présente de la variabilité instantanée que l'on appelle aussi « volatilité » et que celle-ci soit dépendante du temps. Celle-ci peut être constante comme elle peut être très instable. Un exemple représentatif possédant une

volatilité changeante est le prix de l'essence au Québec. Le modèle ARCH (*Autoregressive Conditional Heteroskedasticity*) fut proposé par Engle [Eng82] pour représenter une telle variabilité instantanée. Bollerslev [Bol86] généralisa le tout en créant le modèle GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*).

Le nombre de modèles augmente de jour en jour, mais très souvent en se basant sur les premiers modèles créés tels que les AR, MA, ARMA, ARCH et GARCH. Toutefois, chaque modèle possède un objectif commun, celui de prédire la variable d'intérêt en des temps futurs et c'est pourquoi la dépendance dans chaque variable ou chaque série chronologique est un élément très important. Dans un cas multivarié avec N séries chronologiques, la notion de dépendance est double étant donné que l'on peut l'étudier entre les variables d'une même série, mais aussi entre $2, 3, \dots, N$ séries. Étudier cette

double dépendance à des fins de prédiction peut être fort utile.

Abe Sklar [Skl59] introduisit le concept des copules, qui permet maintenant d'étudier la dépendance de manière plus flexible. La flexibilité vient du fait que l'on peut ainsi modéliser et quantifier la dépendance à l'aide d'une distribution de probabilité. La copule permet de décomposer la distribution conjointe en distributions marginales et la structure de dépendance est expliquée par la copule. Malgré l'apparition du concept en 1959, ce n'est qu'à la fin du 20^e siècle que l'étude et l'utilisation des copules commencent à prendre de l'ampleur. On peut penser notamment aux publications de Deheuvels [Deh80], Frees et Valdez [FV98], Joe [Joe97] et Nelsen [Nel99]. Ces travaux ouvrirent la porte aux copules dans de nombreux domaines tels que les statistiques, les sciences actuarielles et la finance. Le début des années 2000 a vu une émergence impressionnante d'articles et de livres, comme Genest et Remillard [GR04], Patton [Pat09] ou Chen et Fan [CF06], pour ne citer que quelques exemples.

La dépendance chez une copule paramétrique se trouve dans son ou ses paramètre(s).

L'atout que possède les copules est que ce paramètre ne sert qu'à mesurer la dépendance. Les fonctions marginales peuvent être étudiées séparément, ce qui se révèle être une très grande qualité de la modélisation à l'aide de copules. Dans la littérature, les chercheurs se sont surtout penchés sur le cas des séries chronologiques stationnaires.

Alors, la dépendance recherchée est fort probablement constante, et ne devrait pas augmenter ou diminuer avec le temps. Chercher une valeur constante pour cette dépendance n'est pas approprié pour toutes les séries chronologiques, notamment les non-stationnaires. Ce mémoire se veut un pas dans cette direction, par l'analyse de séries chronologiques non-stationnaires, présentant une dépendance qui change dans le temps. Nous analyserons ainsi le cas d'un processus markovien d'ordre 1, à l'aide de copules bivariées à paramètre dynamique.

Au chapitre 1, nous aborderons d'abord quelques modèles classiques pour les séries

chronologiques. Par la suite, au chapitre 2, nous introduirons le concept des copules ainsi que deux familles de copules intéressantes, et nous discuterons de certaines mesures de dépendance très utiles, ainsi que de méthodes d'estimation pour obtenir le paramètre de la copule. Au chapitre 3, nous réviserons la littérature en ce qui concerne les copules à paramètre constant, et discuterons des copules à paramètre dynamique dans le cas des séries chronologiques univariées. Le chapitre 4 offrira des applications théoriques et empiriques.

CHAPITRE 1

Les séries chronologiques

Les statisticiens étudient souvent un phénomène pour mieux le comprendre ou le prédire, ou les deux. Ces phénomènes que l'on peut analyser à partir de données peuvent être étudiés de trois façons : la coupe transversale, les séries longitudinales et les séries temporelles ou chronologiques. Par contre, l'intérêt de connaître l'influence du temps sur des données nous amène à regarder la troisième façon de plus près, soit l'analyse chronologique. Une façon de prédire une variable aléatoire est de s'intéresser à la dépendance que possède cette variable dans le temps. Bien que plusieurs méthodes existent afin de modéliser cette dépendance, l'utilisation de la copule est très intéressante dans ce contexte. Permettant de représenter la dépendance temporelle entre une variable aléatoire et elle-même à des temps différents ou entre deux variables aléatoires au même temps ou à des temps différents par une fonction de répartition, la copule est un outil très flexible. On peut utiliser un certain type de copule lorsque la dépendance dans les données est forte aux premiers temps de la série et plus faible, pour un même écart, vers la fin, et choisir un autre type de copule dans une situation différente. Ainsi, les copules une fois appliquées aux séries chronologiques deviennent un

atout important pour la description et la prédiction de nombreux phénomènes.

Une série chronologique ou temporelle est une suite de données observées à intervalles de temps successifs de longueur égale. Au fil du temps, de nombreux statisticiens inventèrent des modèles dédiés à l'étude des séries chronologiques. Par contre, quelques généralités s'imposent avant de décrire certains de ces modèles ou certaines familles de modèles.

Une série chronologique peut présenter quatre types d'effets : l'effet de tendance, l'effet cyclique, l'effet saisonnier et l'effet irrégulier. On identifie l'effet de tendance à la croissance ou la décroissance sur un long intervalle de temps. L'effet cyclique est quant à lui représenté par des fluctuations régulières ou périodiques autour de la tendance de notre série s'il y en a une, ou autour du niveau moyen de la série sinon. L'effet saisonnier est un patron dans le graphique des données qui se répète sur une certaine période de temps. Enfin, l'effet irrégulier représente le tracé de ce qui reste de la série une fois les effets de tendance, cyclique et saisonnier enlevés.

Le but d'étudier ces effets pour une série est d'arriver à mieux comprendre de quoi sont faites les données. Un objectif souvent poursuivi est de rendre la série stationnaire. Par définition, une série stationnaire est une série chronologique n'ayant aucun effet de tendance à long terme, que ce soit de la croissance ou de la décroissance. On peut penser par exemple à l'évolution de la population au Canada. Sauf catastrophe majeure (guerre ou pandémie), cette série devrait croître à long terme, la population étant grandissante, et la condition de stationnarité ne serait pas respectée. De plus, une série (faiblement) stationnaire est aussi homoscedastique, c'est-à-dire que la dispersion des données est constante. Il existe en effet plusieurs types de stationnarité. Avant de les définir, nous devons introduire la fonction d'autocovariance ainsi que la fonction d'autocorrélation.

Dans une série temporelle, on peut se questionner à savoir si la donnée obtenue au temps t est dépendante d'une donnée précédente. Par exemple, les ventes d'une entreprise en janvier 2019 (temps t) dépendent ou sont très souvent liées à celles de décembre 2018 (temps $t - 1$). Cependant, les observations faites au mois de janvier 2019 pourraient ressembler à celles faites au mois de janvier 2018, d'où une dépendance entre le temps t et le temps $(t - 12)$ est possible et peut être importante. C'est pourquoi une fonction représentant la dépendance entre une variable au temps t et la même variable au temps $(t - h)$ devient très utile. Celle-ci se nomme la fonction d'autocovariance

$\{\gamma(h)\}_{h \in \mathbb{Z}}$ de délai h et s'exprime comme suit, pour une série $(Y_t)_{t \in \mathbb{Z}}$:

$$\gamma(h) = \mathbb{E}[(Y_t - \mathbb{E}(Y_t))(Y_{t-h} - \mathbb{E}(Y_{t-h}))] = Cov(Y_t, Y_{t-h}) \quad (1.1)$$

où $\mathbb{E}(\cdot)$ représente l'opérateur d'espérance. Pour un processus stationnaire, la fonction d'autocovariance est une fonction paire et est semi-définie positive.

La fonction d'autocorrélation $\{\rho(h)\}_{h \in \mathbb{Z}}$, comme son nom l'indique, mesure quant à elle la corrélation entre une variable et elle-même à des temps différents, par exemple avec un délai de h unités de temps. On la définit par :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z}. \quad (1.2)$$

Ainsi, la fonction d'autocorrélation $\rho(h)$ exprime la force et le signe du lien entre Y_t et Y_{t-h} , qui transite par les variables entre celles-ci $(Y_{t-1}, \dots, Y_{t-h+1})$ si $|h| \geq 2$. Par contre, il peut être intéressant d'étudier le lien entre Y_t et Y_{t-h} , sans que les variables entre ces deux dernières ne soient incluses. La fonction d'autocorrélation partielle mesure un tel lien et est définie comme suit :

$$\phi_{hh} = Corr(Y_t - P(Y_t|Y_{t-1}, \dots, Y_{t-h+1}), Y_{t-h} - P(Y_{t-h}|Y_{t-1}, \dots, Y_{t-h+1})).$$

où $P(Y_t|Y_{t-1}, \dots, Y_{t-h+1})$ représente la régression de Y_t sur l'ensemble $\{Y_{t-1}, \dots, Y_{t-h+1}\}$ et $h > 1$. De plus, $\phi_{11} = \rho(1)$ par définition. L'égalité ci-dessus peut

se ré-écrire comme un ratio de deux déterminants :

$$\phi_{hh} = \frac{|P_h^*|}{|P_h|} \quad (1.3)$$

avec

$$P_h^* = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(h-2) & \rho(1) \\ \rho(1) & 1 & \dots & \rho(h-3) & \rho(2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \dots & \rho(1) & \rho(h) \end{pmatrix}$$

et

$$P_h = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(h-2) & \rho(h-1) \\ \rho(1) & 1 & \dots & \rho(h-3) & \rho(h-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho(h-1) & \rho(h-2) & \dots & \rho(1) & 1 \end{pmatrix}$$

L'utilité de ces fonctions d'autocorrélation (acf) et d'autocorrélation partielle (pacf) sera de pouvoir déterminer l'ordre du modèle autorégressif ainsi que celui du modèle de moyenne mobile.

Maintenant que ces fonctions sont définies, nous pouvons désormais introduire le concept de stationnarité mentionné précédemment. D'abord, la stationnarité d'une série chronologique (Y_1, Y_2, \dots, Y_t) , au sens faible, est par définition :

- $\mathbb{E}(Y_t) = \mu, \forall t = 1, \dots, T$
- $Var(Y_t) = \sigma^2 < \infty, \forall t = 1, \dots, T$
- $Cov(Y_t, Y_{t-h}) = \gamma(h), \forall t = 1, \dots, T, \forall h = 1, \dots, T.$

On remarque que le temps n'a aucune influence ici. Effectivement, l'espérance et la variance de la série restent constantes, peu importe la valeur de t . De plus, si la position dans le temps n'a aucun impact, alors la troisième condition est remplie. De son côté, la stationnarité forte est définie quant à elle par :

- Pour toute fonction f mesurable, $f(Y_1, Y_2, \dots, Y_t)$ et $f(Y_{1+h}, Y_{2+h}, \dots, Y_{t+h})$ possèdent la même loi.

Ainsi, on s'intéresse à avoir une fonction de densité jointe semblable à travers la série temporelle. De plus, comme la loi de probabilité d'une distribution d'une série chronologique est difficile à estimer, on se retrouve très souvent à utiliser la stationnarité au sens faible.

La dépendance à travers une série chronologique est quelque chose de logique à vouloir étudier. De plus, lorsqu'on possède plusieurs séries temporelles, on peut s'intéresser à l'effet de causalité entre ces séries. Cette causalité est celle définie au sens de Granger [Gra69]. Dans un cas où l'on aurait deux séries temporelles $(X_t)_{t \in \mathbb{Z}}$ et $(Y_t)_{t \in \mathbb{Z}}$, cette causalité s'obtient en s'intéressant à prédire Y_t à l'aide de X_t . Une autre forme de causalité pour une seule série chronologique est de s'intéresser à l'influence d'une même variable à des temps distincts. Cette causalité nécessite l'utilisation du bruit blanc. Afin d'introduire cette notion, il nous est impératif d'abord d'introduire celle du bruit blanc.

Par définition, le bruit blanc est un processus faiblement stationnaire ayant certaines propriétés supplémentaires. En effet, l'espérance d'un bruit blanc doit être nulle, et la variance se doit d'être finie. La fonction d'autocorrélation d'un bruit blanc est d'ailleurs nulle, indiquant que chaque temps d'un bruit blanc est non-corrélé l'un envers l'autre.

Un bruit blanc $(\epsilon_t)_{t \in \mathbb{Z}}$ respecte donc les conditions suivantes :

- $\mathbb{E}(\epsilon_t) = 0$,
- $\mathbb{E}(\epsilon_t^2) < \infty$,

— $\rho(k) = 0, \forall k \geq 1$.

On note d'ailleurs $\epsilon_k \sim N(0, \sigma_\epsilon^2)$ lorsqu'un bruit blanc est indépendant et identiquement distribué de loi normale, qui porte le nom de bruit blanc gaussien. Même si le bruit blanc ne porte pas une structure de corrélation triviale, et qu'il est un processus purement aléatoire, il joue un rôle très important dans la modélisation d'un processus.

Maintenant que nous savons ce qu'est un bruit blanc, nous pouvons aborder le concept de causalité. Ici, la causalité représente le lien à travers les variables d'une même série chronologique. Si un processus $(X_t)_{t \in \mathbb{Z}}$ peut s'écrire comme une combinaison linéaire de bruit blanc à des temps $t - k$, avec $k \in \mathbb{N}$ et $\sum_{k=0}^{\infty} \alpha_k < \infty$, alors ce processus est

causal :

$$X_t = \sum_{k=0}^{\infty} \alpha_k \epsilon_{t-k}. \quad (1.4)$$

Si à l'inverse, on peut écrire un bruit blanc comme une combinaison linéaire des observations du processus, nous aurons alors un processus dit inversible. Ainsi, $(X_t)_{t \in \mathbb{Z}}$ est inversible si et seulement s'il existe $\beta_k \in \mathbb{R}$ avec $k \in \mathbb{N}$ et $\sum_{k=0}^{\infty} \beta_k < \infty$ tels que

$$\epsilon_t = \sum_{k=0}^{\infty} \beta_k X_{t-k}. \quad (1.5)$$

Obtenir la stationnarité pour notre série possède le grand avantage d'utiliser les propriétés de la stationnarité. Or, pour la rendre stationnaire, nous devons enlever les quatre types d'effets nommés précédemment, s'ils sont présents dans la série. L'effet irrégulier peut être représenté par des événements tels que des grèves ou des attaques comme celle du 11 septembre 2001. Prévoir ces phénomènes n'est pas possible, et c'est ce que l'on considérera comme un bruit blanc. L'effet cyclique peut être éliminé à l'aide d'un déphasage, et les effets de tendance et de saisonnalité peuvent tous simplement être estimés à l'aide de méthodes paramétriques ou de méthode non-paramétriques tels que le lissage exponentiel ou le filtre moyenne mobile.

Pour simplifier l'écriture, un opérateur de délais peut être utile. Pour un processus stochastique $(X_t)_{t \in \mathbb{Z}}$ et un opérateur de délais B , nous avons $BX_t = X_{t-1}$. Dans le même ordre d'idée, un opérateur de différence d'ordre d , où d représente l'horizon, noté ∇_d , nous permet d'obtenir $\nabla_d X_t = X_t - X_{t-d}$. Ces deux opérateurs nous permettent de simplifier l'écriture, mais peuvent aussi être utilisés afin d'éliminer la composante saisonnière.

Comme mentionné précédemment, l'étude des séries chronologiques est très utile et très populaire afin d'obtenir une bonne qualité d'ajustement et d'arriver à obtenir de bonnes prédictions. L'utilisation de modèle afin d'arriver à prédire les valeurs d'une série ou simplement d'arriver à reproduire une série temporelle devient un atout non négligeable. Dans la sous-section qui suit, les modèles les plus utilisés seront abordés. Nous discuterons aussi de certains modèles permettant de régler le cas où la série chronologique initiale ne serait pas stationnaire.

1.1 Les types de modèles

Les différents modèles existants pour la modélisation statistique sont nombreux. J'introduirai les plus populaires adaptés au domaine des séries chronologiques dans cette section. Ceux-ci sont les modèles ARMA et (G)ARCH. Je discuterai aussi en bref des modèles ARIMA et ARFIMA. Le modèle ARMA étant une combinaison du modèle autorégressif (AR) et du modèle de moyenne mobile (MA), il sera donc nécessaire d'introduire ceux-ci avant d'aborder le modèle ARMA. Le modèle ARMA trouvera son utilité lorsqu'un modèle AR ou MA nécessitera un nombre élevé de paramètres à estimer. On pourra ainsi réduire le nombre d'estimations à effectuer et obtenir une meilleure précision. Le modèle autorégressif à hétéroscédasticité conditionnelle (ARCH) permettra de détecter la volatilité dans une série chronologique, c'est-à-dire le

changement dans la variance. Le modèle (G)ARCH représente le modèle ARCH généralisé. Plus particulièrement, le modèle (G)ARCH reconnaît la différence entre un modèle à variance conditionnelle et inconditionnelle, et permet donc d'obtenir une variance non constante à travers le temps grâce à une fonction des erreurs passées.

1.2 Processus à moyenne mobile (MA)

La suite $(X_t)_{t \in \mathbb{Z}}$ est un processus à moyenne mobile d'ordre q ($q \in \mathbb{N}_+$) si on peut l'écrire sous la forme :

$$X_t = \mu + \sum_{k=0}^q \theta_k \epsilon_{t-k} + \epsilon_k \quad (1.6)$$

où $\mu \in \mathbb{R}$, $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ et les θ_k pour $k = 1, 2, \dots, q$ sont les paramètres du modèle, avec $\theta_q \neq 0$. Ainsi, on peut noter $X_t \sim \text{MA}(q)$.

On peut remarquer qu'un processus $\text{MA}(q)$ est, par définition, un processus causal. De plus, on peut exprimer un processus à moyenne mobile à l'aide de l'opérateur de retard :

$$X_t = \mu + \theta(B)\epsilon_t \quad (1.7)$$

avec

$$\theta(B) = \theta_0 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q = \sum_{k=0}^q \theta_k B^k \quad (1.8)$$

où $\theta_0 = 1$ et $\theta_q \neq 0$. Pour un processus $\text{MA}(q)$, lorsque X_t est inversible, c'est-à-dire que

$\theta(z) \neq 0$, $z \in \mathbb{C}$ avec $|z| \leq 1$, on obtient :

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[\mu + \theta(B)\epsilon_t] \\ &= \mu + \theta(B)\mathbb{E}[\epsilon_t] \\ &= \mu \end{aligned}$$

et

$$\begin{aligned}
Var(X_t) &= Var(\mu + \theta(B)\epsilon_t) \\
&= (1^2 + \theta_1^2 + \dots + \theta_q^2) Var(\epsilon_t) \\
&= \sigma_\epsilon^2 \sum_{k=0}^q \theta_k^2.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\gamma(h) &= Cov(X_t, X_{t-h}) \\
&= Cov(\mu + \theta(B)\epsilon_t, \mu + \theta(B)\epsilon_{t-h}) \\
&= \begin{cases} Var(X_t) = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\epsilon^2 & \text{si } h = 0 \\ (\theta_h + \theta_1\theta_{h+1} + \dots + \theta_q\theta_{q-h}) \sigma_\epsilon^2 & \text{si } 0 < h \leq q \\ 0 & \text{si } h > q. \end{cases}
\end{aligned}$$

Ceci signifie que

$$\rho(h) = \begin{cases} 1 & \text{si } h = 0 \\ \frac{(\theta_h + \theta_1\theta_{h+1} + \dots + \theta_q\theta_{q-h})}{(1 + \theta_1^2 + \dots + \theta_q^2)} & \text{si } 0 < h \leq q \\ 0 & \text{si } h > q. \end{cases}$$

La variance et l'espérance n'étant pas dépendantes du temps, les processus MA(q) sont stationnaires pour toutes valeurs de $q > 0$. De plus, le résultat pour notre fonction d'autocorrélation est très intéressant sur le plan des simulations. Effectivement, pour toute valeur $h > q$, on obtient une autocorrélation de 0, mais surtout, pour $h = q$, on obtient $\rho(q) = \theta_q / (1 + \theta_1^2 + \dots + \theta_q^2)$ avec $\theta_q \neq 0$. Ainsi, pour une série chronologique quelconque, chercher à partir de quel ordre les autocorrélations seront nulles nous permet de trouver le modèle moyenne mobile d'ordre q optimal. Ainsi, pour la série temporelle de la figure 1.1, représentant un processus MA(4) de paramètres $\theta_k = (0.8, -0.5, 0.3, 0.2)$ et $\sigma_\epsilon^2 = 1$, on remarque qu'après la quatrième autocorrélation, les suivantes ne dépassent pas un seuil pour être considérées non-nulles. On obtient alors un processus moyenne mobile d'ordre 4.

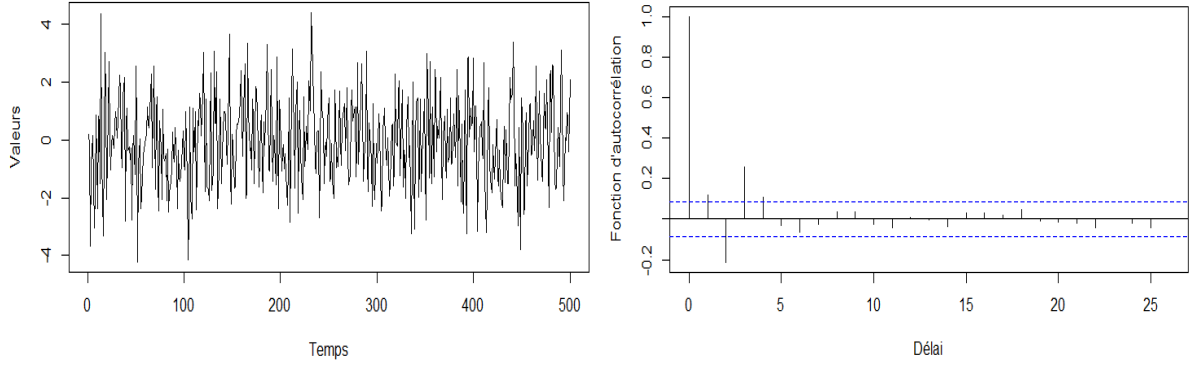


Figure 1.1 – Graphique d’un processus MA(4) de 500 observations avec paramètres $\theta_1 = 0.8$, $\theta_2 = -0.5$, $\theta_3 = 0.3$, $\theta_4 = 0.2$, $\sigma_\epsilon^2 = 1$, $\mu = 0$ et de sa fonction d’autocorrélation

1.3 Processus autorégressif (AR)

La suite $(X_t)_{t \in \mathbb{Z}}$ est un processus autorégressif d’ordre p ($p > 0$) si on peut l’écrire sous la forme :

$$X_t = \nu + \sum_{k=1}^p \phi_k X_{t-k} + \epsilon_t \quad (1.9)$$

où $\nu \in \mathbb{R}$, $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$ et les ϕ_k pour $k = 1, 2, \dots, p$ sont les paramètres du modèle, avec $\phi_p \neq 0$. Ainsi, on peut noter $X_t \sim \text{AR}(p)$. L’opérateur de retard nous permet de récrire notre processus :

$$X_t = \nu + \phi(B)X_t + \epsilon_t \quad (1.10)$$

avec

$$\phi(B) = \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p = \sum_{k=1}^p \phi_k B^k \quad (1.11)$$

Par convention, $\phi_0 = 1$ et $\phi_p \neq 0$.

Un processus $\text{AR}(p)$ est causal et stationnaire si et seulement si son polynôme $\phi(z)$ est

tel que :

$$\phi(z) \neq 0 \text{ avec } z \in \mathbb{C} \text{ tel que } |z| \leq 1$$

Autrement dit, toutes les valeurs des racines du polynôme $\phi(z)$ sont à l'extérieur du cercle complexe unitaire. (voir démonstration dans Brockwell et Davis [BD02])

Pour un processus AR(p) tel que $\mathbb{E}(X_t) = \mathbb{E}(X_{t-h})$, pour $h = 0, 1, \dots, p$, on obtient :

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[\nu + \phi(B)X_t + \epsilon_t] \\ &= \nu + \mathbb{E}[\phi(B)X_t] \\ &= \nu + \mathbb{E}[X_t] \sum_{k=1}^p \phi_k, \end{aligned}$$

ce qui implique

$$\mathbb{E}[X_t] = \frac{\nu}{1 - \sum_{k=1}^p \phi_k}.$$

On peut obtenir la variance et la fonction d'autocorrélation suivante :

$$\begin{cases} \text{Var}(X_t) &= \phi_1\gamma(1) + \phi_2\gamma(2) + \dots + \phi_p\gamma(p) + \sigma_\epsilon^2 \\ \rho(h) &= \begin{cases} 1 & \text{si } h = 0 \\ \sum_{k=1}^{h-p+1} \phi_k \rho(h-k) & \text{si } h \in \mathbb{Z}^*. \end{cases} \end{cases}$$

L'équation de Yule-Walker peut être utilisée afin d'estimer les paramètres du modèle.

Elle est donnée par

$$\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \dots + \phi_p\gamma(h-p) = \sum_{k=1}^p \phi_k\gamma(h-k).$$

Cette équation s'obtient en multipliant (X_{t-h}) des deux côtés de l'égalité à l'équation

(1.9) et en y appliquant l'espérance par la suite.

Il suffit ensuite de prendre l'équation ci-haute et de l'inverser, soit d'exprimer les coefficients en fonction des autocovariances, pour estimer les paramètres.

Avec cette structure, les processus autorégressifs sont tous considérés comme inversibles, peu importe leur ordre. D'ailleurs, il est intéressant de savoir qu'il est possible de passer d'un processus MA(1) à un processus AR(∞), si $|\theta| < 1$, de la façon suivante :

$$\begin{aligned} X_t &= \epsilon_t - \theta \epsilon_t \\ &= (1 - \theta B) \epsilon_t. \end{aligned}$$

Alors,

$$\begin{aligned} \epsilon_t &= \frac{X_t}{(1 - \theta B)} \\ &= X_t + \theta B X_t + \theta^2 B^2 X_t + \dots \\ &= X_t + \sum_{k=1}^{\infty} \theta_k X_{t-k}. \end{aligned}$$

Par conséquent,

$$X_t = - \sum_{k=1}^{\infty} \theta_k X_{t-k} + \epsilon_t.$$

Par la même technique, on peut aussi passer d'un processus AR(1) à un processus MA(∞) si $|\phi| < 1$:

$$X_t = \phi X_{t-1} + \epsilon_t,$$

donc

$$\begin{aligned} \epsilon_t &= X_t - \phi X_{t-1} \\ &= (1 - \phi B) X_t. \end{aligned}$$

Ceci implique

$$\begin{aligned} X_t &= \frac{\epsilon_t}{(1 - \phi B)} \\ &= \epsilon_t + \phi B \epsilon_t + \phi^2 B^2 \epsilon_t + \dots \\ &= \epsilon_t + \sum_{k=1}^{\infty} \phi^k \epsilon_{t-k}. \end{aligned}$$

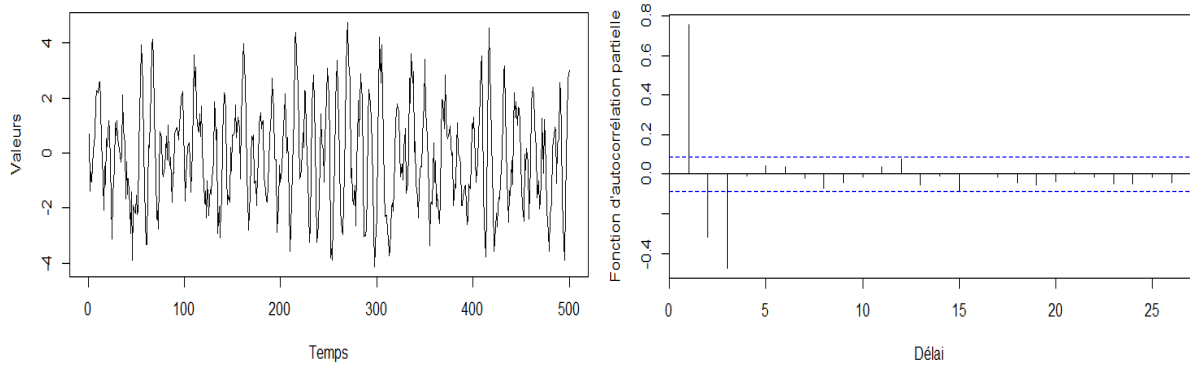


Figure 1.2 – Graphique d’un processus AR(3) de 500 observations avec paramètres $\phi_1 = 0.8$, $\phi_2 = 0.2$, $\phi_3 = -0.5$, $\sigma_\epsilon^2 = 1$, $\nu = 0$ et de sa fonction d’autocorrélation partielle

Tout comme les modèles MA(q), on peut déterminer quel modèle AR(p) est le mieux adapté pour une série chronologique. On applique la même procédure, mais cette fois-ci, sur les autocorrélations partielles. Par exemple, l’ordre p du modèle autorégressif sera la dernière autocorrélation plus grande que le seuil représenté par les lignes pointillées. Le

modèle AR(3) de la figure 1.2 avec comme paramètre $\phi_k = (0.8, 0.2, -0.5)$, $\sigma_\epsilon^2 = 1$ et $\nu = 0$ nous donne les autocorrélations partielles du graphique de droite. On remarque qu’effectivement, l’ordre du processus autorégressif trouvé serait de $p = 3$, étant donné que les autocorrélations partielles à partir du quatrième délai valent 0.

1.4 Processus autorégressif et moyenne mobile (ARMA)

Le processus ARMA(p, q) est une généralisation des modèles AR(p) et MA(q). Le modèle ARMA permet d’obtenir généralement des approximations de séries temporelles avec moins de paramètres qu’un modèle AR ou MA. Pour une suite $(X_t)_{t \in \mathbb{Z}}$, un

processus est dit ARMA d'ordre (p, q) s'il satisfait

$$X_t + \sum_{k=1}^p \phi_k X_{t-k} = \mu + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}$$

avec $\mu \in \mathbb{R}$, $\{\epsilon\} \sim N(0, \sigma_\epsilon^2)$ et $\{\phi_k, \theta_j\}$, $k = 1, \dots, p$, $j = 1, \dots, q$ les paramètres du modèle. Par convention, $\phi_p \neq 0$ et $\theta_q \neq 0$. On peut noter à l'aide de l'opérateur de retard le modèle comme :

$$\phi(B)X_t = \delta + \theta(B)\epsilon_t$$

avec $\phi(B)$ et $\theta(B)$ définis par (1.8) et (1.11).

Par définition, un modèle $\text{ARMA}(p, 0) = \text{AR}(p)$ et un modèle $\text{ARMA}(0, q) = \text{MA}(q)$. Ces égalités représentent bien la flexibilité du modèle ARMA, ce dernier pouvant être représenté par nos deux modèles vus précédemment. Pour qu'un modèle ARMA soit stationnaire, les racines de $\phi(z)$ doivent être à l'extérieur du cercle complexe unitaire. Pour qu'un modèle ARMA soit inversible, les racines de $\theta(z)$ doivent être à l'extérieur du cercle complexe unitaire. En plus de ces deux conditions, pour qu'un modèle ARMA soit à la fois stationnaire et inversible, il ne faut pas que $\phi(z)$ et $\theta(z)$ possèdent des racines communes.

Afin de trouver l'ordre p et q d'un modèle ARMA, nous devons d'abord nous assurer que notre série est stationnaire. Si la série ne semble pas stationnaire, des techniques pour enlever la tendance ou la saisonnalité peuvent être utilisées. Une piste à prendre peut être de différencier la série, ce que nous étudierons à la Section 1.6. Une fois la stationnarité obtenue, on peut déterminer l'ordre q à l'aide de la fonction d'autocorrélation et l'ordre p avec la fonction d'autocorrélation partielle. La méthode pour obtenir ces paramètres a d'abord été élaborée par Box et Jenkins en 1976 [BJ76], avant d'être améliorée par la suite par ces mêmes auteurs en 1990 [BJ90] ainsi que Brockwell et Davis [BD91] en 1991.

Pour un processus ARMA stationnaire, on obtient :

$$\mathbb{E}(X_t) = \frac{\mu}{1 - \sum_{k=1}^p \phi_k}.$$

De plus, la fonction d'autocorrélation d'un processus ARMA(p, q) est donnée par :

$$\rho(h) - \phi_1 \rho(h-1) - \dots - \phi_p \rho(h-p) = 0 \text{ pour } h > q$$

L'exemple illustré à la figure 1.3 représente un processus ARMA(3, 1). On remarque que les fonctions d'autocorrélation et d'autocorrélation partielles données par la figure 1.4 ne nous permet pas d'estimer les ordres p et q . Effectivement, les acf et pacf de certaines séries temporelles peuvent nous offrir des résultats visuels plus ou moins clairs. Valider le modèle à l'aide de critères tels que le critère d'information d'Akaike (AIC) [Aka74] ou du critère d'information bayésien (BIC) [Sch78] de Schwarz devient fort intéressant. Ces deux critères peuvent aussi être utilisés dans nos modèles présentés précédemment. Utiliser le critère du AIC ou du BIC signifie simplement que nos paramètres estimés seront ceux qui maximiseront la vraisemblance et qui minimiseront ces équations :

$$AIC = -2 \log(\mathcal{L}) + 2(p + q) \tag{1.12}$$

$$BIC = -2 \log(\mathcal{L}) + (p + q) \log T \tag{1.13}$$

où T représente le nombre d'observations et $\log(\mathcal{L})$ la log-vraisemblance de notre modèle.

1.5 Processus autorégressif conditionnellement hétéroscédastique (généralisé)

Jusqu'aux années 1980, les modèles ARMA furent les plus utilisés dans le domaine des séries chronologiques. Par contre, une particularité de ces modèles est que le processus

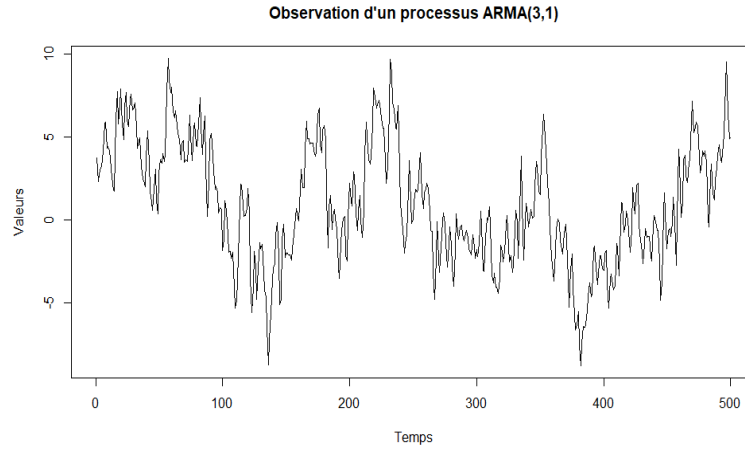


Figure 1.3 – Graphique d'un processus ARMA(3,1) de 500 observations avec paramètres $\phi_1 = 0.8$, $\phi_2 = -0.2$, $\phi_3 = 0.3$, $\theta_1 = 0.8$, $\sigma_\epsilon^2 = 1$ et $\mu = 0$

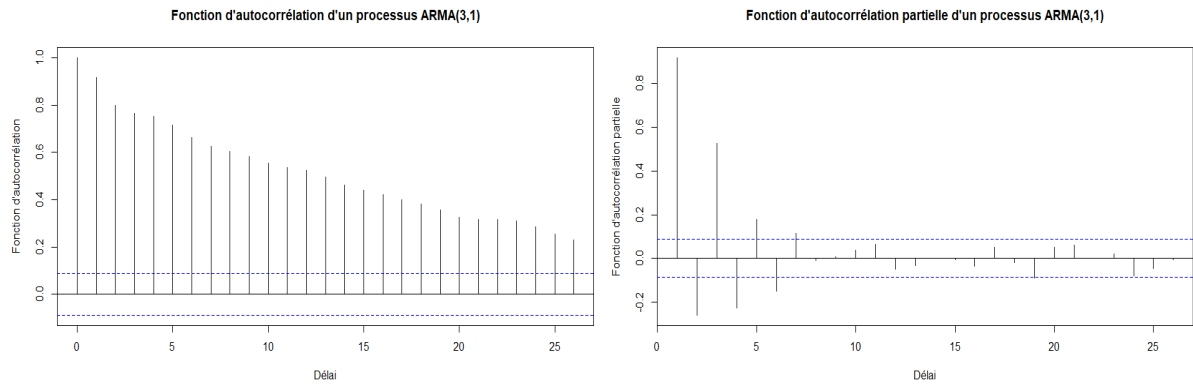


Figure 1.4 – Graphiques des fonctions d'autocorrélation et d'autocorrélation partielle associées au processus ARMA(3,1) de la figure 1.3

$(X_t)_{t \in \mathbb{Z}}$ possède une variance constante, déterminée par les bruits blancs (ϵ_t) . Ainsi, ces modèles ne sont pas adéquats pour une série dont la volatilité change à travers le temps. Les modèles ARCH introduits par Engle (1982) sont une solution à ce problème. Définis par une paramétrisation endogène de la variance conditionnelle, les modèles ARCH(q) sont représentés par :

$$X_t = \sigma_t \epsilon_t \quad (1.14)$$

et

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j X_{t-j}^2 \quad (1.15)$$

avec comme contraintes $\alpha_0 > 0$ et $\alpha_j \geq 0$ pour $j = 1 \dots q$. De plus, ϵ_t est un processus i.i.d et indépendant du passé de X_t . On observe donc que

$$\mathbb{E}(X_t | X_{t-1}, \dots, X_0) = 0$$

et que

$$Var(X_t | X_{t-1}, \dots, X_0) = \sigma_t^2.$$

Simplement en regardant le graphique de la série chronologique étudiée, on peut déterminer si un modèle ARCH peut être utile ou non pour bien réussir à estimer la série. Il s'agit de regarder si la variance semble constante ou non. Dans l'exemple de la figure 1.5, représentant un processus ARCH(2) avec paramètre $\alpha_0 = 0$, $\alpha_1 = 0.2$ et $\alpha_2 = 0.4$, on peut remarquer une grande variation des valeurs. Ainsi, l'utilisation d'un modèle ARCH serait utile. L'utilisation d'une méthode telle que celle des moindres carrés pondérés peut être utile dans l'estimation des paramètres.

Une généralisation naturelle de ce processus est le modèle GARCH. Ce dernier permet de restreindre le nombre de paramètres à estimer, mais aussi d'offrir une solution pour

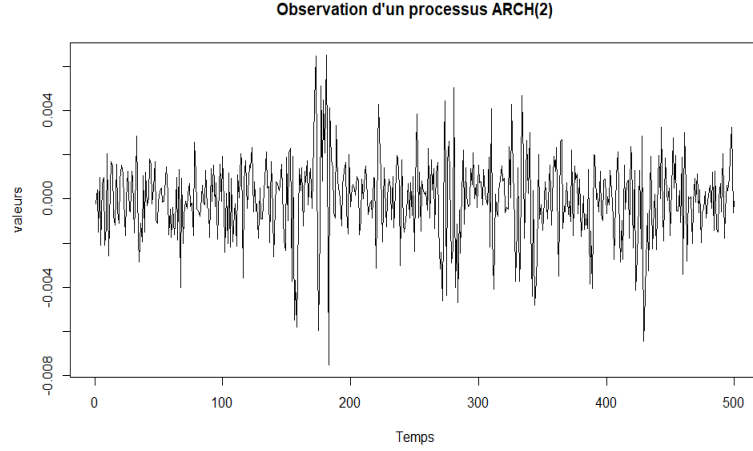


Figure 1.5 – Graphique d'un processus ARCH(2) de 500 observations avec paramètres $\alpha_0 = 0$, $\alpha_1 = 0.2$, $\alpha_2 = 0.4$

les processus à mémoire longue. On définit un modèle GARCH(p, q) appliqué à une série $(X_t)_{t \in \mathbb{Z}}$ comme

$$X_t = \sigma_t \epsilon_t \quad (1.16)$$

et

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j X_{t-j}^2 + \sum_{k=1}^p \beta_k \sigma_{t-k}^2 \quad (1.17)$$

avec comme contraintes $\alpha_0 > 0$, $\alpha_j \geq 0$ et $\beta_k \geq 0$ pour $j = 1 \dots q$, $k = 1, \dots, p$. Tout comme pour le modèle ARCH, ϵ_t est un processus i.i.d et indépendant du passé de X_t .

Le graphique de la figure 1.6 représente un processus GARCH(1, 1) avec paramètres $\alpha_0 = 0$, $\alpha_1 = 0.2$ et $\beta_1 = 0.8$. Tout comme pour le graphique du modèle ARCH, on remarque facilement à l'oeil que nos observations possèdent une variance dépendante du temps.

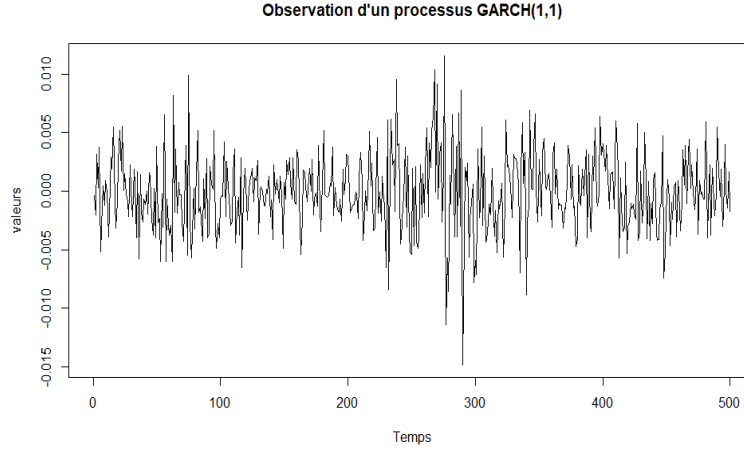


Figure 1.6 – Graphique d’un processus GARCH(1, 1) de 500 observations avec paramètres $\alpha_0 = 0$, $\alpha_1 = 0.1$, $\beta_1 = 0.8$

1.6 Processus autorégressif et moyenne mobile intégré (ARIMA)

La méthodologie utilisée jusqu’à présent pour les modèles $AR(p)$, $MA(q)$ et $ARMA(p, q)$ permet de décrire le comportement d’une série temporelle stationnaire.

Ainsi, pour ces modèles, il est impératif de vérifier si la série est stationnaire avant d’estimer les ordres. En présence d’une série non stationnaire, une technique qui permet de résoudre les problèmes de tendance est la transformation proposée par Box-Jenkins (1976) [BJ76]. Cette transformation se nomme la différenciation et est associée à la lettre I (*autoregressive integrated moving average*) dans ARIMA. La différenciation correspond à l’approximation de la dérivée de la série.

Par définition, pour une série chronologique (X_1, X_2, \dots, X_T) , la différenciation d’ordre 1 prend la forme suivante :

$$Y_t = (1 - B)X_t \quad (1.18)$$

pour $t = 2, 3, \dots, T$.

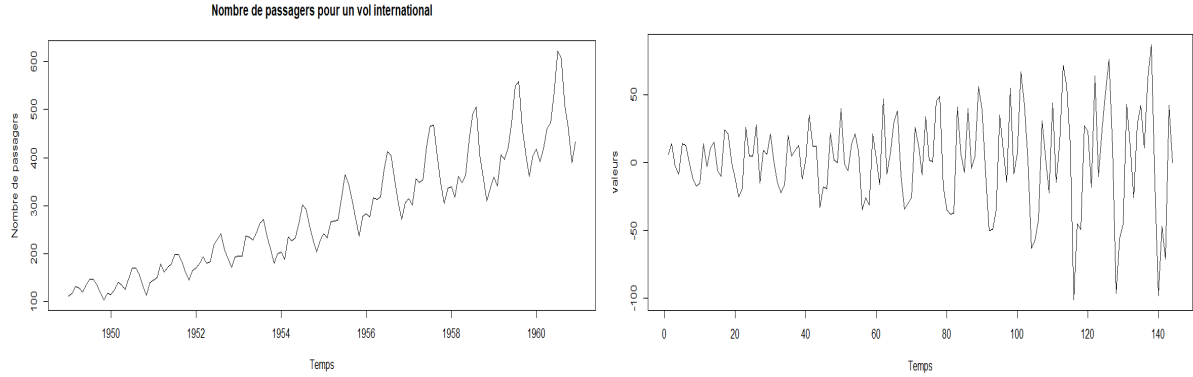


Figure 1.7 – Graphiques du nombre de passagers par vol international à gauche et de sa différentiation d’ordre 1 à droite

Si le problème n’est pas réglé après une différentiation d’ordre 1, appliquer de nouveau une différentiation sur la série différenciée nous permet d’obtenir l’ordre 2. Il suffit donc

d’appliquer des différentiations jusqu’à l’obtention de la stationnarité. Un modèle $ARIMA(p, d, q)$ est représenté par un processus $AR(p)$ défini en (1.9), un processus $MA(q)$ défini en (1.6) et une différentiation d’ordre d . On obtient le modèle

$Y_t = X_t - X_{t-d}$. Si aucune différentiation n’est requise pour rendre le modèle stationnaire ($d = 0$), alors le modèle à optimiser est un $ARMA(p, q)$. Voici par exemple une série non stationnaire représenté par le graphique de gauche à la figure 1.7, qui à

l’aide d’une différentiation d’ordre 1, devient la série stationnaire du graphique de droite. Cependant, on remarque qu’après différentiation d’ordre 1, une grande volatilité est observée et l’application d’un modèle (G)ARCH à cette nouvelle série deviendrait intéressante. Un modèle ARMA-GARCH pourrait aussi être utile afin de résoudre ce problème, celui-ci étant défini comme

$$Y_t = aX_t + \epsilon_t$$

$$\text{où } \epsilon_t \sim GARCH(p, q).$$

1.7 Processus autorégressif et moyenne mobile intégré fractionnaire (ARFIMA)

À première vue, comme une série temporelle est stationnaire ou non, l'utilisation des modèles ARMA ou ARIMA règle tous les problèmes et nous offre les solutions nécessaires pour l'obtention de nos objectifs. Par contre, le modèle ARMA renvoie un processus à mémoire courte, étant donné l'absence de racine unitaire ($d = 0$), et le modèle ARIMA renvoie un processus à mémoire infinie, avec une racine unitaire ($d = 1$). L'étude d'un cas intermédiaire représentant un processus à mémoire longue, représenté par des valeurs de d fractionnaire, pourrait être le meilleur choix afin d'estimer nos paramètres p et q . Ce modèle est le ARFIMA(p, d, q), avec d un réel au lieu d'un entier. Ce modèle n'étant pas l'objet de l'étude, je vous invite si vous souhaitez en savoir plus sur le sujet à consulter l'ouvrage de Hosking (1981) [Hos81].

CHAPITRE 2

Les copules

En statistique, une copule permet de décrire la structure de dépendance entre deux variables aléatoires ou plus. Peu utilisé avant 1950, ce concept est en pleine expansion dans tous les domaines où la dépendance entre variables est utile. Dans ce chapitre, nous débuterons par quelques propriétés d'une copule, pour continuer avec les différentes familles paramétriques, ainsi que des méthodes d'estimation de paramètres.

2.1 Définitions et propriétés

Par définition, la copule nous permet d'expliquer une fonction de répartition multivariée à l'aide de la fonction de répartition univariée de chacune des variables. La copule est donc elle-même une fonction de répartition, que l'on notera C , définie sur $[0, 1]^d$ avec des marges uniformes sur $[0, 1]$.

Ainsi, pour $H(x_1, x_2, \dots, x_d)$ une fonction de répartition d-variée et F_i la fonction de

répartition univariée de X_i pour $i = 1, \dots, d$, on obtient l'égalité suivante :

$$H(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)). \quad (2.1)$$

C'est ce que l'on appelle le théorème de Sklar (1959). De plus, si toutes les lois marginales F_i sont continues, la copule C est unique et on peut l'obtenir à partir de (2.1) :

$$C(x_1, x_2, \dots, x_d) = H(F_1^{-1}(x_1), F_2^{-1}(x_2), \dots, F_d^{-1}(x_d)). \quad (2.2)$$

C'est Sklar qui introduisit la notion de copule, du mot latin *copula*, signifiant « coupler », en lien avec les lois marginales. Son théorème met en lumière le fait que la loi H s'exprime par $(d + 1)$ éléments, soient les d lois marginales F_i , $i = 1, \dots, d$, ainsi que la copule C . La dépendance des variables se retrouve ainsi seulement dans C .

Une copule est une fonction $C : [0, 1]^d \rightarrow [0, 1]$ qui possèdent trois propriétés caractéristiques. Notons $u_i = F_i(x_i)$. Ces propriétés sont :

1. Pour tout $u_i \in [0, 1]$, $C(u_1, u_2, \dots, u_d) = 0$ si une des composantes $u_i = 0$.
2. $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$
3. La fonction C est d-croissante.

Pour $d = 2$, la troisième condition ci-dessus s'exprime comme suit, pour tous

$$u_1, u_2, v_1, v_2 \in [0, 1] \text{ tels que } u_1 \leq u_2 \text{ et } v_1 \leq v_2 :$$

$$C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) \geq 0.$$

La copule étant une fonction de répartition multivariée, il est donc possible d'obtenir la densité de la copule, notée $c(u_1, u_2, \dots, u_d)$:

$$c(u_1, u_2, \dots, u_d) = \frac{\partial^d C(u_1, u_2, \dots, u_d)}{\partial u_1 \partial u_2 \dots \partial u_d}. \quad (2.3)$$

Tout comme il est possible de passer d'une loi marginale multivariée H à une fonction copule C composée de fonctions de répartition F_i ($i = 1, \dots, d$), on peut exprimer la densité de la distribution conjointe notée h à l'aide de la densité de copule c , des fonctions de répartition F_i et des fonctions de densité marginale f_i de X_i , avec $i = 1, \dots, d$. On obtient ainsi l'égalité suivante :

$$h(x_1, x_2, \dots, x_d) = f_1(x_1) \cdot f_2(x_2) \cdot \dots \cdot f_d(x_d) \cdot c(F_1(x_1), F_2(x_2), \dots, F_d(x_d)). \quad (2.4)$$

L'équation 2.4 est très intéressante parce que les paramètres de la copule seront estimés grâce à celle-ci.

La fonction copule possède plusieurs avantages par rapport aux indicateurs de dépendance tels que le coefficient de Pearson, le rho de Spearman et le tau de Kendall, pour définir la dépendance entre plusieurs variables. Ses principales utilités sont sa flexibilité et sa description complète de la structure de dépendance. Commençons par définir les indicateurs de dépendances précédemment nommés.

2.2 Mesures de dépendance

La mesure la plus connue est la corrélation linéaire entre deux variables, habituellement notée $\frac{Cov(X, Y)}{\sigma_X \sigma_Y}$. Cette corrélation permet de quantifier à quel point la relation linéaire est forte entre plusieurs variables aléatoires continues. Soit une paire de vecteurs (X, Y) représentant deux séries chronologiques (x_1, x_2, \dots, x_n) et (y_1, y_2, \dots, y_n) telles que n est la longueur des séries, alors le coefficient de corrélation de Pearson est donné par

$$r(X, Y) = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\left(\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \right) \left(\sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} \right)} = \frac{\widehat{Cov}(X, Y)}{s_X s_Y} \quad (2.5)$$

où s_X et s_Y représentent les écarts-types empiriques et $\widehat{Cov}(X, Y)$ représente la covariance empirique des échantillons que constituent les deux séries. De plus, $r \in [-1, 1]$. Une valeur de 1 signifie que lorsque X augmente, Y augmente et que les observations représentent une droite croissante. Une valeur de -1 signifie que lorsque X augmente, Y diminue, et donc que la droite est décroissante. Une valeur de 0 signifie qu'il n'y a aucun lien linéaire à attendre entre X et Y.

Une autre mesure de corrélation sur les rangs est donnée par $\frac{Cov(rg(X), rg(Y))}{\sigma_{rg(X)}\sigma_{rg(Y)}}$. Cette mesure de dépendance entre deux variables aléatoires est non-paramétrique. On utilise habituellement cette mesure lorsqu'on a des doutes à propos de la normalité de la distribution des variables et que l'on présume que les variables ne sont pas nécessairement corrélées linéairement. Contrairement au coefficient de corrélation de Pearson, le calcul de cette corrélation ne se fait pas par rapport aux valeurs du vecteur (X, Y) , mais par rapport aux rangs de celles-ci. Le rho de Spearman est donné par

$$\begin{aligned}\rho(X, Y) &= \frac{\frac{1}{n-1} \sum_{i=1}^n (rg(x_i) - rg(\bar{x})) (rg(y_i) - rg(\bar{y}))}{\left(\sqrt{\frac{1}{n-1} \sum_{i=1}^n (rg(x_i) - rg(\bar{x}))^2} \right) \left(\sqrt{\frac{1}{n-1} \sum_{i=1}^n (rg(y_i) - rg(\bar{y}))^2} \right)} \\ &= \frac{\widehat{Cov}(rg(X), rg(Y))}{s_{rg(X)} s_{rg(Y)}}\end{aligned}\tag{2.6}$$

où $s_{rg(X)}$ et $s_{rg(Y)}$ représentent les écarts-types empiriques des rangs des deux échantillons ou séries et $\widehat{Cov}(rg(X), rg(Y))$ représente la covariance empirique entre les deux vecteurs de rangs. Comme pour le coefficient de corrélation de Pearson, les valeurs de rho sont entre -1 et 1. Par contre, une valeur de 1 ou -1 signifie dans le cas de Spearman que la dépendance entre les variables X et Y est monotone parfaite. Les valeurs du rho de Spearman seront positives si la tendance est croissante, et négative lorsque décroissante. Une valeur de 0 pour ρ signifie ici qu'il n'y a pas de tendance monotone entre X et Y. On peut écrire le rho de Spearman en fonction de la copule C

comme suit :

$$\rho(C) = -3 + 12 \int_{[0,1]^2} C(u, v) du dv \quad (2.7)$$

où $u = F_1(X_1)$ et $v = F_2(X_2)$.

La troisième et dernière mesure abordée dans ce texte est le tau de Kendall. Elle mesure l'association entre deux variables à l'aide de la concordance. Comme le rho de Spearman, elle mesure la dépendance à l'aide du rang. Pour un ensemble d'observations (X, Y) avec $X = (x_1, x_2, \dots, x_n)$ et $Y = (y_1, y_2, \dots, y_n)$, les paires (x_i, y_i) et (x_j, y_j) ($i \neq j$) sont dites concordantes si $x_i < x_j$ et $y_i < y_j$ ou $x_i > x_j$ et $y_i > y_j$. Dans un cas contraire, on dira que les paires sont discordantes. Si $x_i = x_j$ et $y_i = y_j$, la paire ne sera ni discordante, ni concordante. Le calcul pour obtenir le tau de Kendall empirique est le

suivant :

$$\tau_n(X, Y) = \frac{(\text{nombre de paires concordantes}) - (\text{nombre de paires discordantes})}{\frac{1}{2} \cdot n \cdot (n - 1)}. \quad (2.8)$$

Le dénominateur représente le nombre total de paires possible. Cette mesure prend également ses valeurs entre -1 et 1 .

Un atout important du tau de Kendall est qu'on peut l'écrire en fonction de la fonction copule (Schweizer et Wolff 1981) [SW81]. Comme nous nous intéresserons par la suite à des processus markoviens d'ordre 1 que nous représenterons comme un couple

(X_t, X_{t-1}) , seuls les résultats pour le cas bivarié sont montrés :

$$\begin{aligned} \tau(C) &= 4 \cdot \mathbb{E}[C(F_1(X_1), F_2(X_2))] - 1 \\ &= 1 - 4 \int_0^1 \int_0^1 \left(\frac{\partial C(u_1, u_2)}{\partial u_1} \right) \left(\frac{\partial C(u_1, u_2)}{\partial u_2} \right) du_1 du_2. \end{aligned} \quad (2.9)$$

Par contre, un défaut du tau de Kendall que possèdent aussi le rho de Spearman et le coefficient de corrélation de Pearson est la limitation de leur flexibilité. Que ce soit pour

calculer la dépendance d'un cas bivarié ou multivarié ($d > 2$), le résultat est un chiffre qui tente d'expliquer globalement la dépendance des variables étudiées. Utiliser une mesure où la dépendance peut varier selon les variables sélectionnées apporterait la flexibilité recherchée. C'est exactement ce que fait la copule, en séparant la marginale d -variée en d marginales indépendantes. Par contre, il faut faire attention à la copule choisie. Plusieurs classes et familles de copules existent, toutes ayant des propriétés différentes, de sorte que la copule représente un choix crucial. De plus, chaque copule possède un paramètre de dépendance. Trouver la meilleure copule pour une série implique de sélectionner la famille de copule adéquate et d'estimer le meilleur paramètre pour cette famille. Tel est le sujet des prochaines sections.

2.3 Classes et familles de copules

Même si les copules sont apparues dans les années 1950, ce n'est que vers les années 1980-90 que la recherche sur le sujet a avancé grandement. L'utilité des copules étant de modéliser la dépendance de façon plus flexible, plusieurs familles et classes de copule ont été développées et proposées. Dans cette section, nous nous intéresserons d'abord aux classes, puis aux familles à l'intérieur de celles-ci. Les copules étudiées seront la gaussienne, la Student, la Gumbel, la Clayton, et la Frank.

2.3.1 Classe elliptique

La classe elliptique tire son nom de ce que généralement, les courbes de niveau de la densité possèdent une forme elliptique. Par définition, une copule est dite elliptique si l'égalité suivante est vérifiée :

$$c(u) = (\det \Sigma)^{-1/2} g((u - \mu) \Sigma^{-1} (u - \mu)^\top) \quad (2.10)$$

où Σ est une matrice carrée de dimension $d \times d$ et définie positive, $\mu = (\mu_1, \mu_2, \dots, \mu_d)$ est un vecteur de paramètre de position et \top est l'opérateur transposé. De plus, $g(\cdot)$ doit être une fonction à valeurs positives tel que l'équation $\int_{\mathbb{R}^d} g(uu^\top) du = 1$ est respectée et $\det \Sigma$ désigne le déterminant de Σ .

La copule gaussienne et la copule de Student font partie de cette classe. D'autres copules respectent aussi cette condition pour leur densité de copule, mais seules la copule gaussienne et la Student seront étudiées en détail.

La copule gaussienne

Soient $\Phi(\cdot)$ la fonction de répartition d'une distribution normale centrée et réduite et $\Phi^{-1}(\cdot)$ son inverse. Soient Σ la matrice de variance-covariance et $\Phi_\Sigma(\cdot)$ la fonction de répartition d-variée centrée réduite. La copule gaussienne est représentée par :

$$C_\rho(u_1, u_2, \dots, u_d) = \Phi_\Sigma(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_d)). \quad (2.11)$$

Rappelons que la matrice de variance-covariance est, dans ce cas, une matrice dont l'élément (i, j) représente la corrélation entre les variables u_i et u_j , notée ρ_{ij} . Dans le cas bivarié, la copule peut s'écrire

$$C_\rho(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} f_\rho(s, t) dt ds,$$

où $f_\rho(s, t) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{(s^2 + t^2 - 2\rho st)}{2(1-\rho^2)}\right)$ est la densité normale bivariée centrée réduite avec $\rho \in (-1, 1)$.

La densité de copule, qui découle de la définition des copules elliptiques en (2.10), est

$$c_\rho(u_1, \dots, u_d) = (\det \Sigma)^{-1/2} \cdot \exp\left(-\frac{1}{2} \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_d) \end{pmatrix}^\top (\Sigma^{-1} - I_d) \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_d) \end{pmatrix}\right) \quad (2.12)$$

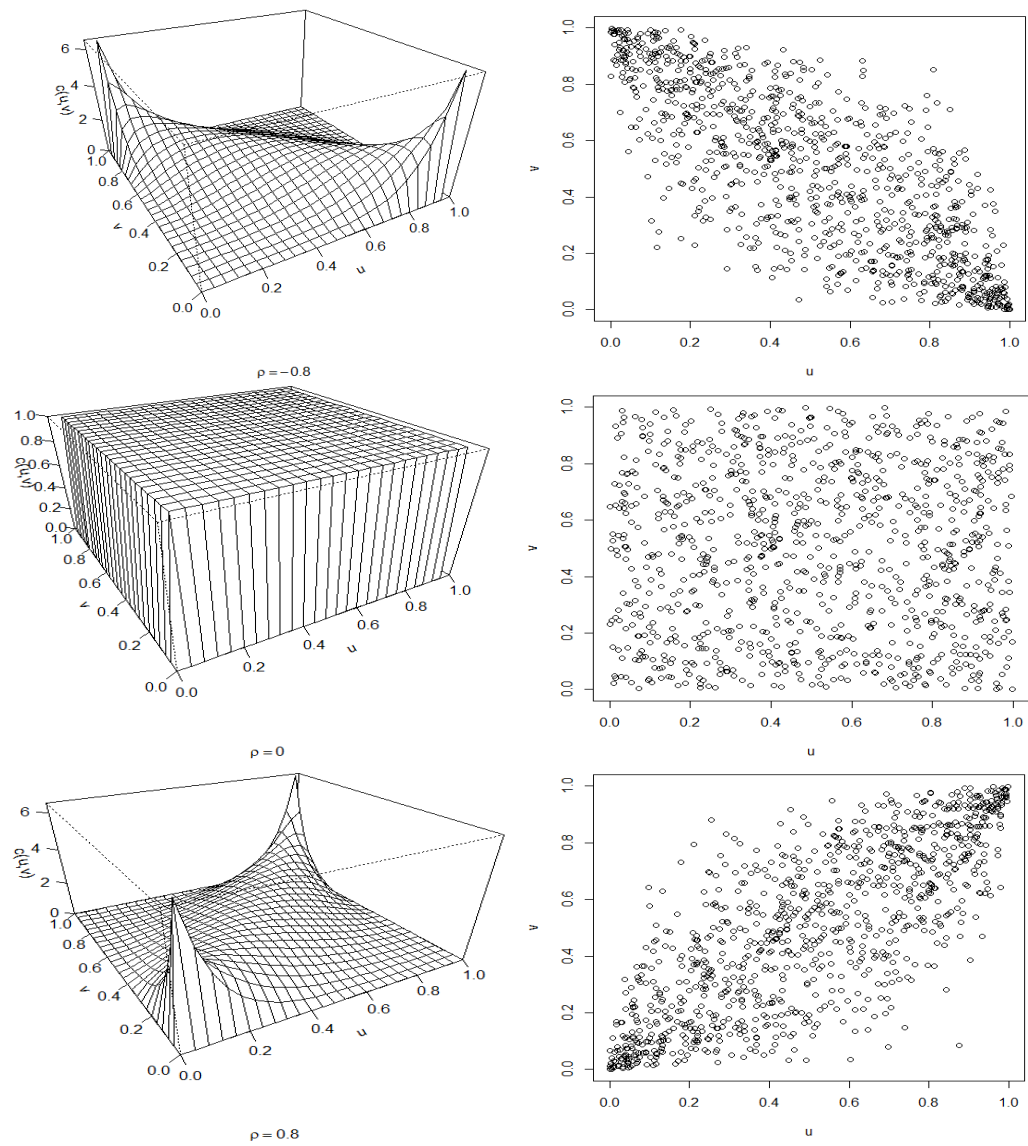


Figure 2.1 – Graphiques de la densité et nuages de points ($n = 1000$) pour des copules gaussiennes bivariées avec paramètre $\rho = -0.8, 0$ et 0.8 .

Notons que quand $\Sigma = I_d$, toute corrélation entre u_i et u_j , $\forall i, j$ tel que $i \neq j$, est nulle et on obtient une copule nommée copule d'indépendance.

Étant donné que la copule gaussienne est issue de la loi normale, cette copule n'est pas très adaptée aux valeurs extrêmes (que ce soit dans la queue de gauche ou de droite).

Le tau de Kendall d'une copule gaussienne obtenu à partir de l'équation (2.9) est

$$\tau = \frac{2}{\pi} \arcsin \rho.$$

La copule de Student

Comme son nom l'indique, cette copule est issue de la distribution multivariée de Student. Celle-ci est définie par

$$C(u_1, u_2, \dots, u_d) = T_{\Sigma, \nu}(T_\nu^{-1}(u_1), T_\nu^{-1}(u_2), \dots, T_\nu^{-1}(u_d)) \quad (2.13)$$

avec T_ν , la fonction de répartition centrée réduite de la distribution de Student, et ν , le nombre de degrés de liberté (ddl) correspondant tel que $\nu \geq 3$. T_ν^{-1} représente l'inverse de la fonction, et $T_{\Sigma, \nu}$, la fonction de distribution multivariée centrée réduite d'une

Student avec un ddl $\nu \geq 3$, où Σ représente, comme dans le cas d'une copule gaussienne, la matrice des corrélations. Dans le cas bivarié, on peut ré-écrire la copule à l'aide d'une double intégrale :

$$C(u, v) = \int_{-\infty}^{T_\nu^{-1}(u)} \int_{-\infty}^{T_\nu^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left(1 + \frac{s^2 + t^2 - 2\rho st}{\nu(1-\rho^2)} \right)^{-(\nu+2)/2} ds dt.$$

La densité de la copule de Student peut quant à elle s'obtenir en appliquant directement

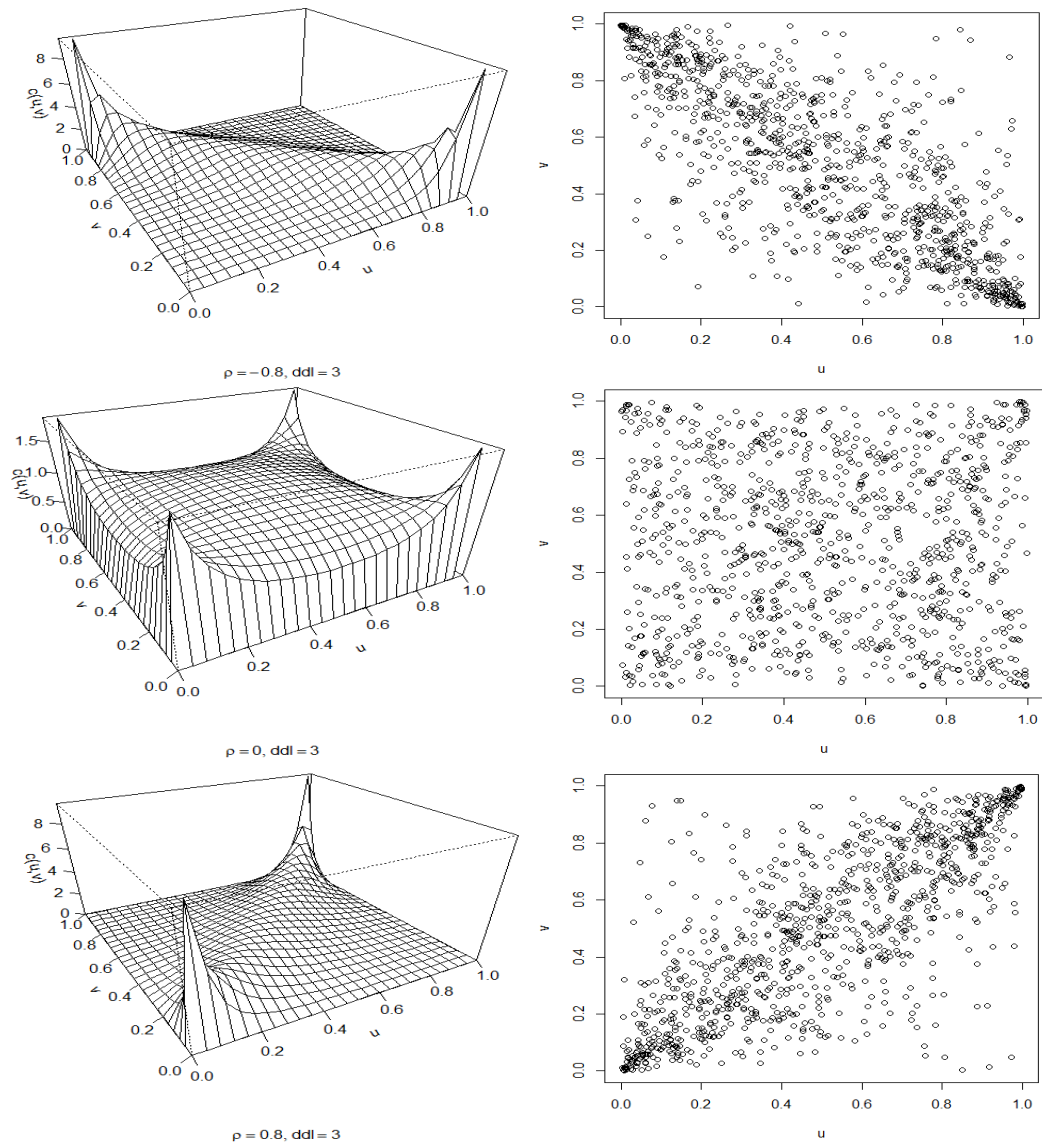


Figure 2.2 – Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Student bivariées avec paramètre $\rho = -0.8, 0, 0.8$ et 3 degrés de liberté (ddl).

la définition d'une densité de copule. Avec $\Lambda = (T_\nu^{-1}(u_1), \dots, T_\nu^{-1}(u_d))^\top$, on obtient

$$c(u_1, \dots, u_d) = |\Sigma|^{-1/2} \frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \left(\frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \right)^d \frac{\left(1 + \frac{1}{\nu} \Lambda^\top \Sigma^{-1} \Lambda\right)^{-\frac{\nu+d}{2}}}{\prod_{i=1}^d \left(1 + \frac{(T_\nu^{-1}(u_i))^2}{\nu}\right)^{-\frac{\nu+d}{2}}},$$

où $\Gamma(\cdot)$ représente la fonction gamma. Même si les ressemblances sont notables entre la copule de Student et la copule gaussienne, la différence majeure est que la copule de Student arrive à capter la dépendance aux extrêmes (dépendance de queue). La formule du tau de Kendall $\frac{2}{\pi} \arcsin(\rho)$ tient aussi pour les lois de Student avec $\nu \geq 3$ et même les lois elliptiques avec densités absolument continues par rapport à la mesure de Lebesgue dont la matrice de variance-covariance existe. [FFK02]

Dépendance de queue

Le concept de dépendance de queue peut s'avérer intéressant pour les copules quand il s'agit de capturer la dépendance dans les quadrants inférieur gauche et supérieur droit d'une distribution bivariable. Afin d'arriver à évaluer cette dépendance, il est nécessaire d'introduire l'inverse des fonctions marginales ainsi que le quasi-inverse pour celles qui ne sont pas strictement croissantes. Le quasi-inverse de F est défini par

$$F^{-1}(b) = \inf\{x | F(x) \geq b\}.$$

De plus, si X_1 et X_2 sont deux variables aléatoires possédant des fonctions de répartition univariées F_1 et F_2 et si les limites existent, la dépendance de queue à droite est définie par

$$\lambda_U = \lim_{b \rightarrow 1^-} \mathbb{P}(X_2 > F_2^{-1}(b) | X_1 > F_1^{-1}(b))$$

et la dépendance de queue à gauche est définie par

$$\lambda_L = \lim_{b \rightarrow 0^+} \mathbb{P}(X_2 \leq F_2^{-1}(b) | X_1 \leq F_1^{-1}(b)) .$$

Le paramètre λ qui exprime la dépendance de queue peut être obtenu à partir de la fonction copule, à droite

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - C(u, u)}{1 - u}$$

et à gauche

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} .$$

2.3.2 Classe archimédienne

La copule archimédienne est définie par l'égalité suivante :

$$C(u_1, u_2, \dots, u_d) = \begin{cases} \varphi^{-1}(\varphi(u_1), \dots, \varphi(u_d)) & \text{si } \sum_{i=1}^d \varphi(u_i) \leq \varphi(0) \\ 0 & \text{sinon ,} \end{cases}$$

où la fonction $\varphi : [0, 1] \rightarrow [0, \infty)$, nommée générateur, est de classe C^2 , c'est-à-dire que la première dérivée est continue et dérivable et la deuxième dérivée est continue. De plus, $\varphi(1) = 0$, $\varphi'(u) \leq 0$ et $\varphi''(u) > 0$. Chaque famille de copule faisant partie de la classe archimédienne est définie par une fonction génératrice.

La copule de Gumbel, la copule de Clayton ainsi que la copule de Frank appartiennent à cette classe. Ces trois copules sont étudiées en détail dans la prochaine section.

Contrairement aux copules de la classe elliptique, les copules la classe archimédienne possèdent des paramètres définis sur un domaine différent. Ainsi, on ne peut pas dire qu'une corrélation de 0.8 pour la copule gaussienne équivaut à un paramètre $\theta = 0.8$ pour les copules de la classe archimédienne. Par contre, à l'aide d'une mesure comme le tau de Kendall, il est possible de trouver un paramètre équivalent pour n'importe quelle copule.

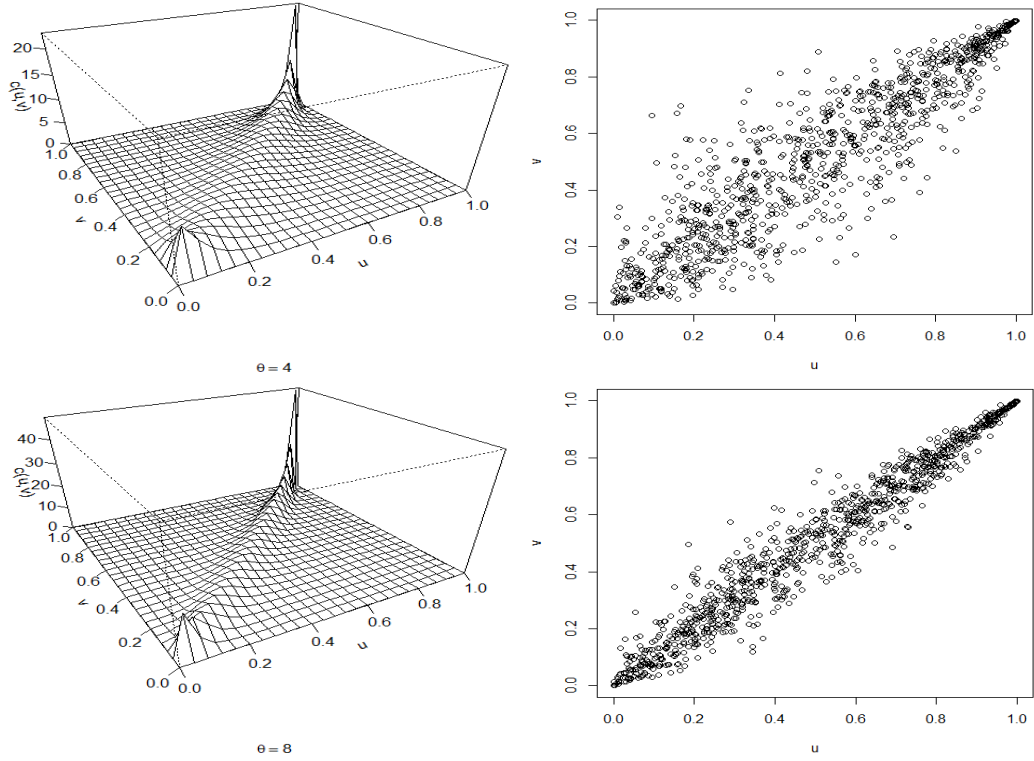


Figure 2.3 – Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Gumbel bivariées avec paramètre $\theta = 4$ et 8 .

La copule de Gumbel

La copule de Gumbel est définie de telle sorte à avoir une dépendance presque nulle dans la queue de gauche, mais une dépendance très forte dans l'extrémité supérieure (queue de droite). Elle est donc très utile pour modéliser la dépendance dans la queue de droite.

Sa fonction génératrice est donnée par $\varphi_\theta(t) = (-\ln t)^\theta$ où $\theta \geq 1$, d'où $\varphi_\theta^{-1}(t) = \exp(-t^{1/\theta})$ et on obtient par définition de la classe archimédienne :

$$C(u_1, u_2, \dots, u_d) = \exp \left(- \left[(-\ln u_1)^\theta + (-\ln u_2)^\theta + \dots + (-\ln u_d)^\theta \right]^{1/\theta} \right). \quad (2.14)$$

En appliquant l'équation (2.9), on obtient que le tau de Kendall pour une copule de

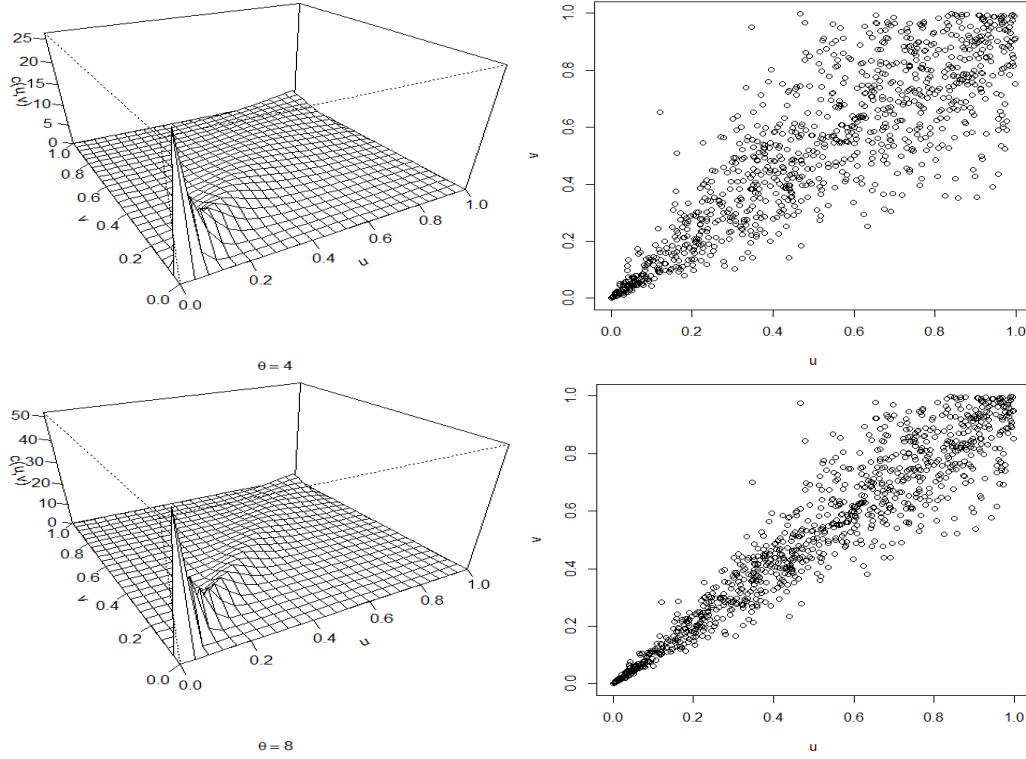


Figure 2.4 – Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Clayton bivariées avec paramètre $\theta = 4$ et 8.

Gumbel est donné par $\tau = 1 - \frac{1}{\theta}$. On peut remarquer que lorsque $\theta = 1$, le tau de Kendall est égal à 0, et lorsque θ tend vers ∞ , τ tend vers 1.

La copule de Clayton

La copule de Clayton est le cas contraire de la copule de Gumbel. La relation entre les fonctions de répartition y est modélisée avec une très grande dépendance dans la queue de gauche, et diminue au fur et à mesure que l'on se déplace vers l'extrémité supérieure.

Cette copule est donc aussi asymétrique.

La fonction génératrice d'une copule de Clayton et son inverse sont données par $\varphi_{\theta}(t) = \frac{1}{\theta}t^{-\theta} - 1$ et $\varphi_{\theta}^{-1}(t) = (\theta t + 1)^{-1/\theta}$ avec $\theta > 0$. On obtient donc comme fonction

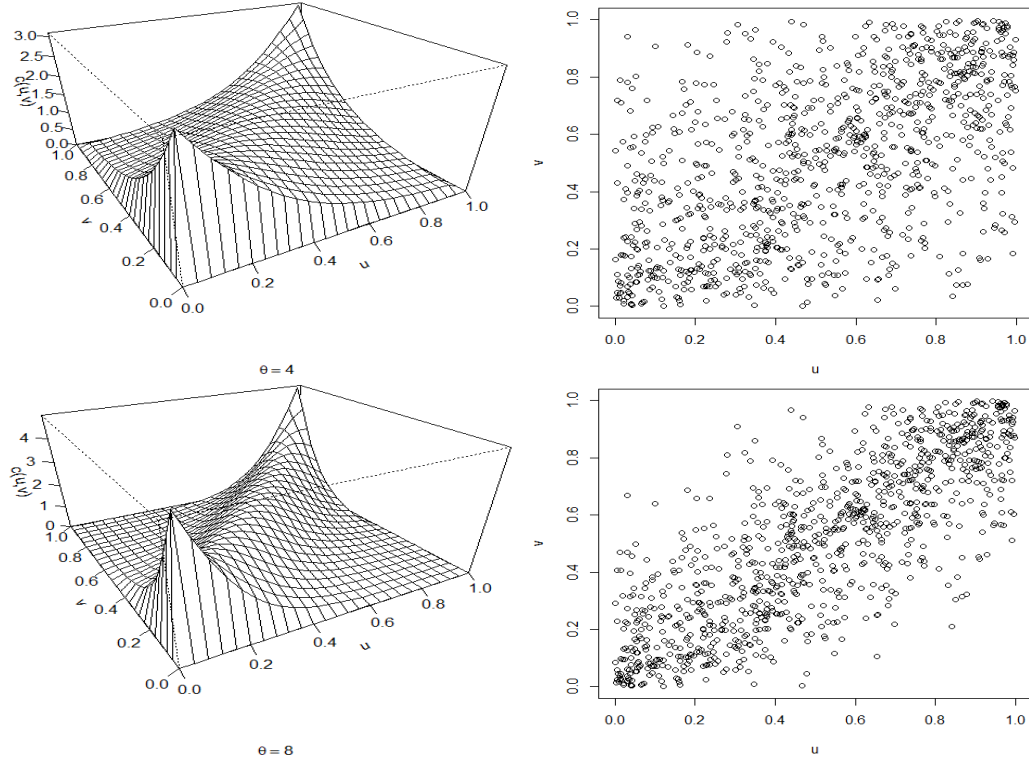


Figure 2.5 – Graphiques de la densité et nuages de points ($n = 1000$) pour des copules de Frank bivariées avec paramètre $\theta = 4$ et 8.

copule :

$$C(u_1, u_2, \dots, u_d) = (u_1^{-\theta}, u_2^{-\theta}, \dots, u_d^{-\theta} - d + 1)^{-\frac{1}{\theta}} \quad (2.15)$$

où $d \geq 2$. Le tau de Kendall d'une copule de Clayton est donné par $\tau = \frac{\theta}{\theta + 2}$.

La copule de Frank

La dernière copule étudiée dans ce mémoire est la copule de Frank. Cette copule est symétrique, c'est-à-dire que la dépendance est la même dans les extrêmes.

Le générateur de cette copule et son pseudo-inverse s'écrivent $\varphi_\theta(t) = -\ln \left(\frac{e^{-\theta t} - 1}{e^\theta - 1} \right)$

et $\varphi_\theta^{-1}(t) = -\frac{1}{\theta} \ln(e^{-t}(e^\theta - 1) + 1)$, avec $\theta \in \mathbb{R} \setminus \{0\}$. Ainsi, la copule de Frank de dimension d est donnée par la fonction copule :

$$C(u_1, u_2, \dots, u_d) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1) \dots (e^{-\theta u_d} - 1)}{(e^{-\theta} - 1)^{d-1}} \right) \quad (2.16)$$

avec $\theta \in \mathbb{R} \setminus \{0\}$. Enfin, le tau de Kendall pour une copule de Frank est $\tau = 1 + \frac{4}{\theta} [D_1(\theta) - 1]$, avec D_1 la fonction de Debye définie par $D_k(x) = \frac{k}{x} \int_0^x \frac{s^k}{e^s - 1} ds$.

2.4 Estimation des copules

Nos copules d'intérêt étant définies, on remarque que chaque famille de copules possède ses avantages et ses limitations. Ainsi, selon les données à l'étude, une copule sortira du lot comme étant celle correspondant le mieux à la dépendance des observations. La copule constitue donc une fonction inconnue en pratique, que l'on cherche à déterminer. Cette section offre différentes méthodes d'estimation de cette fonction : paramétriques, semi-paramétriques, non-paramétriques et par la méthode des moments. De plus, comme la copule est liée aux distributions marginales, celles-ci devront aussi être estimées sauf si connues au préalable.

Estimation paramétrique

Plusieurs méthodes existent afin d'estimer la copule et les marginales de manières paramétriques. Une estimation paramétrique est basée sur un modèle paramétrique utilisé pour estimer les paramètres recherchés, d'où la différence entre les estimateurs proposés se retrouve dans la méthode utilisée. La méthode d'estimation paramétrique la plus connue est sans doute celle proposée par Joe [Joe97], [Joe05] et s'appelle la méthode du maximum de vraisemblance complète (FML : *Full Maximum Likelihood*).

Considérons un cas bivarié où $(x_1, y_1), \dots, (x_n, y_n)$ sont les observations provenant d'une distribution marginale bivariée $H(x, y)$. Ainsi, il existe une fonction copule définie en (2.1) telle que

$$H(x, y) = C(F_1(x, \alpha), F_2(y, \beta); \theta)$$

où θ est le paramètre de la copule et α et β sont les paramètres des marginales.

La méthode FML estime en même temps (α, β, θ) par $(\hat{\alpha}, \hat{\beta}, \hat{\theta})$. Afin d'estimer le paramètre de la copule ainsi que les marginales, on utilise la densité de copule présentée en (2.4) :

$$h(x, y) = c(F_1(x, \alpha), F_2(y, \beta); \theta) \cdot f_1(x) \cdot f_2(y)$$

En découle ensuite la fonction de log-vraisemblance :

$$L(\alpha, \beta, \theta) = \sum_{i=1}^n \log c(F_1(x_i, \alpha), F_2(y_i, \beta); \theta) + \sum_{i=1}^n (\log f_1(x_i, \alpha) + \log f_2(y_i, \beta))$$

Il suffit de maximiser L par rapport au vecteur (α, β, θ) pour obtenir le vecteur d'estimateurs $(\hat{\alpha}, \hat{\beta}, \hat{\theta})$.

Cette méthode d'estimation est très connue et sans doute une des meilleures dans le domaine des fonctions copules. Cependant, ce n'est pas la plus utilisée en raison de la lourdeur des calculs. En effet, cette méthode vise à trouver le meilleur triplet $(\hat{\alpha}, \hat{\beta}, \hat{\theta})$, et procède pour ce faire à toutes les combinaisons possibles de ces trois paramètres.

Rappelons que chaque famille de copules peut, au départ, être la meilleure pour notre échantillon $(x_1, y_1), \dots, (x_n, y_n)$. Le problème s'agit de trouver les meilleurs estimateurs pour chaque famille de copules, et de comparer la vraisemblance de chaque famille afin d'obtenir la *meilleure* copule.

Une méthode tout aussi connue, mais plus utilisée, est la méthode d'inférence pour les marges (IFM : *Inference Functions for Margins*). Celle-ci consiste à suivre la même

démarche que dans la méthode FML, mais en deux étapes. Cette méthode de Joe et Xu [JX96] repose sur l'écriture de la fonction de log-vraisemblance. Les fonctions de densités f_1 et f_2 se trouvant dans la deuxième somme étant séparées du paramètre θ se trouvant dans la première somme, cela permet d'estimer d'abord les densités :

$$\begin{aligned}\hat{\alpha} &= \arg \max \sum_{i=1}^n \log f_1(x_i, \alpha) \\ \hat{\beta} &= \arg \max \sum_{i=1}^n \log f_2(x_i, \beta).\end{aligned}$$

Ceci constitue la première étape. La deuxième étape consiste à estimer le paramètre θ de la copule en utilisant les deux estimateurs obtenus précédemment. Ainsi, il suffit de calculer

$$\hat{\theta} = \arg \max \sum_{i=1}^n \log c(F_1(x_i, \hat{\alpha}), F_2(y_i, \hat{\beta}); \theta).$$

On remarque que la seule et unique différence entre la méthode FML et la méthode IFM est l'estimation en deux étapes, étant donné que la même méthode d'estimation est utilisée dans les deux cas. Cette méthode nécessite beaucoup moins de calculs, car les marginales sont estimées indépendamment du paramètre de la copule. Malgré ce point positif, le paramètre θ est estimé en fonction des estimées des marginales. L'estimateur du paramètre θ sera le *meilleur* par rapport aux marginales estimées, mais le vecteur $(\hat{\alpha}, \hat{\beta}, \hat{\theta})$ obtenu ne sera pas nécessairement optimal.

L'inconvénient que ces deux méthodes ont en commun est le fait que les marginales soient paramétriques. Les paramètres estimés étant associés à des modèles prédéfinis, les résultats obtenus sont très restrictifs par rapport à ceux-ci. Ainsi, l'utilisation d'une approche non-paramétrique, ou une méthode semi-paramétrique représentant un mélange des deux, pourrait améliorer l'estimation des paramètres.

Estimation semi-paramétrique

L'estimation semi-paramétrique signifie par définition que l'estimation des paramètres se fait à l'aide de modèles paramétriques pour les copules, mais fait appel à des modèles non-paramétriques pour les marginales. En fait, la distribution empirique est utilisée pour les marginales, puis comme pour la méthode IFM, le paramètre de la copule est estimé par maximum de vraisemblance. Cette approche porte le nom de méthode d'estimation par maximum de vraisemblance canonique (CML : *Canonical Maximum Likelihood*) qui a été étudiée et proposée par Genest, Ghoudi et Rivest (1995) [GGR95] et Kim (2007) [KSS07].

Les marginales peuvent être estimées à l'aide de la fonction de distribution empirique redimensionnée :

$$\hat{F}_1(x) = \frac{1}{n+1} \sum_{t=1}^n \mathbb{I}(X_t \leq x)$$

et

$$\hat{F}_2(y) = \frac{1}{n+1} \sum_{t=1}^n \mathbb{I}(Y_t \leq y).$$

Par la suite, il ne reste plus qu'à estimer le paramètre θ . En appliquant la même démarche que pour l'étape 2 de la méthode IFM, on obtient l'estimateur

$$\hat{\theta} = \arg \max \sum_{i=1}^n \log c(\hat{F}_1(x_i), \hat{F}_2(y_i); \theta).$$

Pour l'estimation des marginales, la distribution empirique redimensionnée est préférée à la distribution empirique standard. La raison est qu'avec cette distribution, on obtient des marginales strictement inférieures à 1, car la majorité des copules paramétriques ne sont pas bornées en 1 ou 0. En effet, lorsqu'on dérive par rapport à α , β ou θ pour estimer les paramètres, notre fonction de log vraisemblance $\log c(u_1, u_2; \theta)$ vaut l'infini

pour $u_i = 0$ ou 1 ($i = 1, 2$), pour la plupart des densités de copules. Ainsi, pour tout n , la distribution empirique redimensionnée règle ce problème.

Estimation non-paramétrique

L'utilisation d'une méthode non-paramétrique devient intéressante lorsque la spécification d'un modèle paramétrique n'est pas bonne ou facile. Effectivement, l'application d'une méthode paramétrique ou semi-paramétrique peut amener à un mauvais ajustement du modèle aux données.

Pour présenter l'estimation non-paramétrique, il est nécessaire de parler d'abord de la copule empirique, introduite sous le nom de fonction empirique de dépendance par Deheuvels (1980) [Deh80]. Comme le rho de Spearman et le tau de Kendall, la copule empirique se base sur les rangs des observations pour obtenir une structure de dépendance. Pour un échantillon (x_i, y_i) , avec $i = 1, \dots, n$, on estime alors la fonction copule par :

$$\hat{C}(u, v) = \hat{H}(\hat{F}^{-1}(u), \hat{G}^{-1}(v))$$

où $0 \leq u, v \leq 1$ et

$$\begin{aligned}\hat{H}(x, y) &= \frac{1}{n} \sum_{i=1}^n \mathbb{I}(x_i \leq x, y_i \leq y), & -\infty < x, y < \infty \\ \hat{F}(x) = \hat{H}(x, \infty) &= \frac{1}{n} \sum_{i=1}^n \mathbb{I}(x_i \leq x), & -\infty < x < \infty \\ \hat{G}(y) = \hat{H}(\infty, y) &= \frac{1}{n} \sum_{i=1}^n \mathbb{I}(y_i \leq y), & -\infty < y < \infty.\end{aligned}$$

Cet estimateur de copule possède plusieurs propriétés intéressantes, comme par exemple, la convergence asymptotique vers C la vraie copule. Deheuvels a d'ailleurs

établi la convergence faible du processus de copule empirique dans le cas de marges indépendantes.

Un défaut de la copule empirique est qu'elle présente des discontinuités. L'application d'une forme de lissage à cette copule fut une idée adéquate afin de régler ce problème ; voir Fermanian, Radulovic et Wegkamp [FRW02]. Des améliorations telles que la transformation de données ont été proposées ensuite pour cette méthode par Gijbels, Omelka et Veraverbeke [GOV09]. L'utilisation d'un noyau linéaire local à aussi été proposé par Chen et Huang en 2005 [CH05], afin d'éviter le biais aux bords.

Un autre vent d'idée fut de travailler sur la densité de copule au lieu de la fonction copule elle-même ; voir les travaux et résultats de Gijbels et Mielniczuk (1990) [GM90] ainsi que Chen (1999, 2000) [Che99] [Che00], pour ne citer que ceux-ci.

Estimation des paramètres de la copule par la méthodes des moments

Une dernière méthode qui peut être utilisée pour estimer le paramètre de la copule est la méthode des moments. Cette méthode se distingue par sa simplicité d'exécution. Elle requiert l'utilisation du tau de Kendall ou du rho de Spearman, discutés à la Section 2.2. Pour chacune des deux mesures de dépendances, deux méthodes furent proposées pour obtenir un estimateur, soient les équations 2.6 et 2.7 pour le rho de Spearman et les équations 2.8 et 2.9 pour le tau de Kendall. Dans la méthode des moments, le rho de Spearman ou le tau de Kendall, notés $\rho_n(X, Y)$ et $\tau_n(X, Y)$, est estimé d'abord. Il ne reste ensuite qu'à trouver l'estimé $\hat{\theta}$ du paramètre de la copule dans (2.7) à l'aide du rho de Spearman ou (2.9) à l'aide du tau de Kendall, en utilisant l'une des deux égalités suivantes :

$$\begin{aligned}\rho_n(X, Y) = \rho(C(\hat{\theta})) = g(\hat{\theta}) &\implies \hat{\theta} = g^{-1}(\rho_n(X, Y)) \\ \tau_n(X, Y) = \tau(C(\hat{\theta})) = h(\hat{\theta}) &\implies \hat{\theta} = h^{-1}(\tau_n(X, Y)).\end{aligned}$$

Ces égalités peuvent être utilisées quand les distributions marginales F et G pour X et Y sont continues. Dans ce cas, l'estimateur du paramètre de la copule est dit cohérent.

Dans le cas où les distributions marginales sont discrètes, il n'en est plus ainsi et une méthode afin d'estimer le paramètre θ consiste à rendre les marginales continues artificiellement, à l'aide d'un changement de variable pour X et Y (Grebennikov 2014)

[Gre14].

Pour des distributions continues, prenons en exemple la copule de Gumbel avec le tau de Kendall τ comme mesure de dépendance. Il suffit alors d'isoler θ et d'utiliser τ_n comme estimateur de τ . Cette technique porte le nom de tau de Kendall inverse, et les

équations pour chacune des cinq copules étudiées sont données dans le tableau ci-dessous

Copule	Tau de Kendall	Tau de Kendall inverse
Gaussienne	$\tau = \frac{2}{\pi} \arcsin \rho$	$\rho = \sin \left(\frac{\pi}{2} \cdot \tau \right)$
Student		
Gumbel	$\tau = 1 - \frac{1}{\theta}$	$\theta = \frac{1}{1 - \tau}$
Clayton	$\tau = \frac{\theta}{\theta + 2}$	$\theta = \frac{2\tau}{1 - \tau}$
Frank	$\tau = 1 + \frac{4}{\theta} [D_1(\theta) - 1]$	Pas de forme fermée

Tableau 2.1 – Tau de Kendall et tau de Kendall inverse pour chacune des copules d'intérêt, avec $D_k(x) = \frac{k}{x} \int_0^x \frac{s^k}{e^s - 1} ds$

À partir de τ_n obtenu à l'aide de (x_i, y_i) , $i = 1, \dots, n$, on peut construire un estimateur du paramètre de la copule. Un estimateur de celui-ci pour la copule de Gumbel est

donné par

$$\hat{\theta}_{GU} = \frac{1}{1 - \tau_n}.$$

Bien que le rho de Spearman puisse s'exprimer en fonction de la fonction copule, les équations pour lui deviennent rapidement très compliquées (Clayton) ou aucune forme fermée ne peut être obtenue (Gumbel). C'est pourquoi le tau de Kendall est habituellement utilisé pour estimer le paramètre de la copule, nous donnant des équations assez simples, comme on peut le voir dans le tableau ci-dessus.

CHAPITRE 3

Copules statique et dynamique

Dans ce chapitre, nous discuterons d'abord des copules multivariées. La méthode d'estimation privilégiée sera d'estimer le paramètre de la copule de façon paramétrique.

La structure de dépendance sera donc uniquement expliquée par le paramètre de la copule. Cet atout est très important et c'est d'ailleurs ce qui rend les copules si populaires de nos jours. La flexibilité que les copules procurent dans la spécification et l'estimation du modèle permet aux chercheurs de ne pas considérer seulement les distributions multivariées existantes. Par la suite, nous considérerons le cas des copules bivariées et le paramètre de celles-ci deviendra le centre d'intérêt. Le paramètre à valeur constante dans le temps suivi du paramètre variant dans le temps seront analysés.

3.1 La copule multivariée

Tel que vu à la Section 2.4, que ce soit avec un modèle paramétrique ou semi-paramétrique, des estimations pour nos paramètres des marginales et de la copule peuvent être obtenues et un critère tel que l'AIC ou le BIC peut ensuite être utile afin

de déterminer les meilleurs modèles de copule. Dans le cas où plusieurs séries sont observées aux mêmes temps équidistants, le théorème de Sklar ainsi que la notion de la probabilité conditionnelle peuvent nous être utiles dans une procédure inférentielle.

Notons d'abord que

$$f(y|x) = \frac{f_{Y,X}(y, x)}{f_X(x)}.$$

De plus, en appliquant (2.4) pour $v = F_1(y)$ et $u = F_2(x)$, on a que

$$f_{Y,X}(y, x) = c(v, u) \cdot f_1(y) \cdot f_2(x)$$

$$f_X(x) = f_2(x)$$

d'où

$$f(y|x) = \frac{c(v, u) \cdot f_1(y) \cdot f_2(x)}{f_2(x)} = c(v, u) f_1(y) = c(v|u) \cdot f_1(y)$$

pour le cas bivarié où $u \in [0, 1]$.

Dans le cas multivarié, soit y_{it} la i^e série au temps t et $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ le vecteur des n séries étudiées. Posons $\mathbf{Y} \sim \mathbf{H}$ et $Y_i \sim F_i$, avec \mathbf{H} la fonction de distribution jointe. On a alors par (2.1) que

$$\mathbf{H}(\mathbf{y}) = C(F_1(y_1), \dots, F_n(y_n)).$$

Ainsi, peu importe les distributions univariées F_1, \dots, F_n et la distribution multivariée \mathbf{H} , la copule C existe par le théorème de Sklar. La copule enlève donc la restriction que les marginales soient toutes semblables, comme pour la cas d'une gaussienne multivariée. Par exemple, dans un cas bivarié, F_1 pourrait être une distribution d'une loi normale et F_2 pourrait être une distribution d'une loi gamma. La distribution bivariée

obtenue par ces deux distributions ne serait probablement pas une belle équation, mais grâce au théorème de Sklar, l'égalité entre la distribution multivariée et la copule serait valable. Notons \mathcal{F}_{t-1} l'ensemble d'information généré par $\{Y_{t-1}, Y_{t-2}, \dots\}$.

De plus,

$$\begin{aligned} \text{soit } \mathbf{Y}_t | \mathcal{F}_{t-1} &\sim \mathbf{F}(\cdot | \mathcal{F}_{t-1}) \\ \text{et } Y_{it} | \mathcal{F}_{t-1} &\sim F_i(\cdot | \mathcal{F}_{t-1}), \quad i = 1, \dots, n. \end{aligned}$$

Suivant la même démarche que pour le cas bivarié, et en passant de la densité à la fonction marginale, on obtient la version conditionnelle du théorème de Sklar :

$$\mathbf{H}(\mathbf{y} | \mathcal{F}_{t-1}) = C(F_1(y_1 | \mathcal{F}_{t-1}), \dots, F_n(y_n | \mathcal{F}_{t-1}) | \mathcal{F}_{t-1}).$$

Il est important de noter que les marginales et la copule doivent avoir le même ensemble conditionnel, car dans le cas contraire, les marginales conditionnelles pourraient ne pas respecter les conditions d'une distribution conjointe ; voir Fermanian et Wegkamp (2012).

On peut aussi définir les marginales conditionnelles $F_i(Y_{it} | \mathcal{F}_{t-1}) = U_{it}$, ce qui nous permet d'écrire la distribution conditionnelle de $\mathbf{U}_t | \mathcal{F}_{t-1}$ comme la copule conditionnelle de $\mathbf{Y}_t | \mathcal{F}_{t-1}$:

$$\mathbf{U}_t | \mathcal{F}_{t-1} \sim C(\cdot | \mathcal{F}_{t-1}).$$

Ce passage permet de mieux voir l'utilité de la copule pour l'estimation des paramètres.

Effectivement, il suffit d'estimer chacune des U_{it} , représentant les marginales conditionnelles et d'estimer (en même temps ou en deuxième lieu) le paramètre de la copule.

Avant de pouvoir modéliser la dépendance présente entre plusieurs séries, une structure pour le modèle représentant les distributions marginales conditionnelles doit être

déterminée. Celle-ci nous permettra ensuite d'arriver à estimer nos marginales et notre paramètre de copule. Sous le même modèle et les mêmes conditions que Patton (2012)

[Pat12], soit

$$\mathbf{Y}_t = \boldsymbol{\mu}_t(\mathbf{Z}_{t-1}, \boldsymbol{\xi}) + \boldsymbol{\sigma}_t(\mathbf{Z}_{t-1}, \boldsymbol{\xi}) \boldsymbol{\epsilon}_t \quad (3.1)$$

des séries markoviennes où $\boldsymbol{\xi}$ représente le vecteur des paramètres liés aux distributions marginales, $\boldsymbol{\epsilon}_t$ sont les erreurs standardisées, $\mathbf{Z}_{t-1} \in \mathcal{F}_{t-1}$ et

$$\begin{aligned} \boldsymbol{\mu}_t(\cdot, \boldsymbol{\xi}) &= (\mu_{1t}(\cdot, \boldsymbol{\xi}), \mu_{2t}(\cdot, \boldsymbol{\xi}), \dots, \mu_{nt}(\cdot, \boldsymbol{\xi})) \\ \boldsymbol{\sigma}_t(\cdot, \boldsymbol{\xi}) &= \text{diag}\{\sigma_{1t}(\cdot, \boldsymbol{\xi}), \sigma_{2t}(\cdot, \boldsymbol{\xi}), \dots, \sigma_{nt}(\cdot, \boldsymbol{\xi})\} \\ \boldsymbol{\epsilon}_t &= (\epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{nt})^\top \sim F_{\epsilon t} = C_\theta(F_{1t}, F_{2t}, \dots, F_{nt}). \end{aligned}$$

Ici, $\boldsymbol{\mu}_t$ et $\boldsymbol{\sigma}_t$ sont \mathcal{F}_{t-1} -mesurable et ne dépendent pas de $\boldsymbol{\epsilon}_t$. Nous permettrons aux séries d'avoir des moyennes et variances conditionnelles qui changent à travers le temps.

De plus, nous supposons que les erreurs ϵ_{it} auront une distribution conditionnelle constante dans le temps, avec moyenne de 0 et variance de 1.

Avec ces conditions, un modèle optimal pour chacune des séries observées peut être obtenu. Par exemple, le modèle optimal pourrait être un ARMA(1, 1) pour la première série, un AR(2) pour la deuxième, et ainsi de suite, tous des modèles choisis à l'aide d'un critère d'information tel que le AIC ou le BIC. Il sera alors possible d'estimer les paramètres $\boldsymbol{\xi}$ des modèles sélectionnés et de trouver un estimateur des erreurs

standardisées, en isolant $\boldsymbol{\epsilon}_t$ dans (3.1) :

$$\hat{\boldsymbol{\epsilon}}_t = \frac{\mathbf{Y}_t - \boldsymbol{\mu}_t(\mathbf{Z}_{t-1}, \hat{\boldsymbol{\xi}})}{\boldsymbol{\sigma}_t(\mathbf{Z}_{t-1}, \hat{\boldsymbol{\xi}})},$$

où le vecteur $\hat{\boldsymbol{\xi}}$ représente les paramètres estimés du modèle conditionnel à \mathcal{F}_{t-1} .

Par la suite, l'estimation des marginales F_i peut se faire par une méthode paramétrique ou non-paramétrique. Du côté paramétrique, plusieurs choix sont possibles, comme la

distribution normale, la distribution de Student standardisée [Bol86], etc. Du côté non-paramétrique, utiliser la fonction de distribution empirique redimensionnée tel que discuté à la Section 2.4 peut être une solution. Comme la moyenne, la variance ainsi que les erreurs standardisées peuvent être estimées, la distribution conditionnelle est alors totalement connue. Ainsi, une méthode pour estimer les paramètres restants, c'est-à-dire les marginales et le paramètre de la copule, peut être basée sur des simulations. Par contre, il est à noter que la dépendance trouvée sera une fonction de la copule conditionnelle. Ainsi, cette dépendance peut être liée au paramètre de la copule, mais pas nécessairement. De plus, dans le cas où l'on doit estimer les marginales et le paramètre de la copule, chercher à créer un intervalle de confiance pour le paramètre de la copule à l'aide de simulations pourrait nous donner des résultats erronés. Cette erreur résulte du fait qu'une partie de l'erreur se serait produite dans l'estimation des marginales au préalable, à l'aide de la copule conditionnelle. L'importance de bien comprendre l'estimation utilisée est donc cruciale, pour n'importe quel paramètre, notamment les erreurs standardisées.

Revenons au modèle défini en (3.1). Dans les mêmes conditions de travail que Patton (2012) [Pat12], posons la moyenne conditionnelle et la variance conditionnelle du modèle comme étant

$$\begin{aligned}\mathbb{E}[Y_{it}|\mathcal{F}_{t-1}] &= \mu_i(\mathbf{Z}_{t-1}, \boldsymbol{\xi}_i) \text{ pour } \mathbf{Z}_{t-1} \in \mathcal{F}_{t-1} \\ \text{Var}[Y_{it}|\mathcal{F}_{t-1}] &= \sigma_i^2(\mathbf{Z}_{t-1}, \boldsymbol{\xi}_i)\end{aligned}$$

Ces conditions ne sont pas trop restrictives et permettent au modèle ARMA, aux régressions linéaires et non-linéaires, et à bien d'autres de représenter la moyenne conditionnelle adéquatement. La variance conditionnelle peut alors être représentée par un modèle ARCH, ou une extension telle que le GARCH par exemple. Ainsi, la

distribution conditionnelle des erreurs standardisées,

$$\epsilon_{it} = \frac{Y_{it} - \mu_i(\mathbf{Z}_{t-1}, \boldsymbol{\xi}_i)}{\sigma_i(\mathbf{Z}_{t-1}, \boldsymbol{\xi}_i)},$$

peut être traitée de façon paramétrique ou non-paramétrique. On obtient alors

$$\begin{aligned} \epsilon_{it} | \mathcal{F}_{t-1} &\sim F_i(\cdot | \mathbf{Z}_{t-1}, \boldsymbol{\xi}_i) \\ \text{ou } \epsilon_{it} | \mathcal{F}_{t-1} &\sim (i.i.d.) F_i \end{aligned}$$

Ainsi, pour dans le cas non-paramétrique, on suppose que la distribution conditionnelle reste la même, ce qui n'est pas nécessairement vrai dans le cas paramétrique. Cette différence peut permettre au chercheur de favoriser un modèle paramétrique ou non. Le modèle paramétrique permet d'analyser le cas suivant où le paramètre de la copule est constant :

$$\begin{aligned} U_{it} &= F_i(\epsilon_{it}), \quad i = 1, 2, \dots, n \\ \text{ou } \mathbf{U}_t &= (U_{1t}, \dots, U_{nt})^\top | \mathcal{F}_{t-1} \sim (iid) C(\theta^*) \end{aligned}$$

où θ^* est le paramètre de la copule. On cherchera donc à estimer $\delta^* = [\theta^{*\top}, \boldsymbol{\xi}^{*\top}]^\top$ où $\boldsymbol{\xi}^*$ est le paramètre en lien avec les distributions marginales. En utilisant la méthode du maximum de vraisemblance appliquée à la copule conditionnelle avec $\mathbf{f}_t(\cdot | \mathcal{F}_{t-1}; \delta^*)$ la fonction de distribution conditionnelle, on obtient l'estimateur

$$\begin{aligned} \hat{\delta}^* &= \arg \max_{\delta^*} \log L(\delta^*) \\ \text{où } \log L(\delta^*) &= \sum_{t=1}^n \log \mathbf{f}_t(\mathbf{Y}_t | \mathcal{F}_{t-1}; \delta^*) \\ \text{avec } \log \mathbf{f}_t(\mathbf{Y}_t | \mathcal{F}_{t-1}; \delta^*) &= \sum_{i=1}^n f_{it}(Y_{it} | \mathcal{F}_{t-1}; \boldsymbol{\xi}^*) \\ &\quad + \log c(F_{1t}(Y_{1t} | \mathcal{F}_{t-1}; \boldsymbol{\xi}^*), \dots, F_{nt}(Y_{nt} | \mathcal{F}_{t-1}; \boldsymbol{\xi}^*) | \mathcal{F}_{t-1}; \theta^*) \end{aligned}$$

Il ne reste plus que la procédure d'estimation à déterminer. Si l'on veut obtenir tous les estimateurs en une seule étape, on sélectionne la procédure FML, alors qu'en deux étapes, la procédure IFM est suggérée.

Pour la méthode semi-paramétrique, on peut ré-utiliser la procédure CML vue à la Section (2.4), en utilisant la notation ci-dessus. Pour une estimation non-paramétrique des marginales, on peut utiliser

$$\hat{F}_i(\epsilon) = \frac{1}{n+1} \sum_{t=1}^n \mathbb{I}(\hat{\epsilon}_{it} \leq \epsilon)$$

$$\text{où } \hat{\epsilon}_{it} = \frac{Y_{it} - \mu_i(\mathbf{Z}_{t-1}, \hat{\boldsymbol{\xi}})}{\sigma_i(\mathbf{Z}_{t-1}, \hat{\boldsymbol{\xi}})}$$

où $\hat{\boldsymbol{\xi}}$ représente les paramètres estimés du modèle conditionnel à \mathcal{F}_{t-1} . Ainsi, en posant $\hat{U}_{it} = \hat{F}_i(\hat{\epsilon}_{it})$ pour $i = 1, \dots, n$, on obtient l'estimateur du paramètre de la copule comme suit :

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=1}^n \log c_{\theta}(\hat{U}_{1t}, \dots, \hat{U}_{nt}).$$

Une question que l'on peut se poser après l'estimation est : « Quelle confiance peut-on avoir dans les paramètres estimés ». Une réponse peut être apportée par la simulation. Dans la sous-section qui suit, deux algorithmes de simulations dans le cas multivarié sont présentés. Ces algorithmes couvrent le cas où la copule serait de classe elliptique ou archimédienne.

3.1.1 Algorithmes de simulation

Les simulations sont utiles à plus d'un titre. Effectivement, afin de vérifier si une méthode d'estimation est adéquate et offre de bons résultats, il est souvent nécessaire de l'appliquer sur plusieurs séries pour lesquelles on connaît les valeurs des paramètres.

Par exemple, la simulation de T séries caractérisée par un paramètre permet d'obtenir N paramètres estimés différents et ainsi de calculer un écart-type et une moyenne.

L'écart-type fournit une mesure de la précision de l'estimation et la moyenne nous permet de calculer le biais. Il faut garder en tête que chaque classe de copule possède des propriétés uniques. Ainsi, la méthode de simulation se doit d'être différente et appropriée pour chaque classe de copule. Dans cette sous-section, je présenterai deux algorithmes de simulation tirés de Rémillard [RPS12] pour la classe elliptique et pour la classe archimédienne.

Processus markovien et copule elliptique

Par définition, une distribution conjointe est dite meta-elliptique si la copule qui lui est associée est de classe elliptique, c'est-à-dire qu'elle peut être gaussienne ou de Student par exemple. Comme vue en à l'équation 2.10 où ici, $u = \mathbf{Y}$, la densité d'un vecteur \mathbf{Y} est donnée en fonction d'un vecteur de position $\boldsymbol{\mu}$, d'une matrice symétrique définie positive $\boldsymbol{\Sigma}$ et d'un générateur $g(\cdot)$ à valeurs positives. Nous noterons $\mathbf{Y} \sim \mathcal{E}(g, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. De plus, on peut poser $\boldsymbol{\mu} = 0$ sans problème étant donné que la copule de \mathbf{Y} ne dépend que de $\boldsymbol{\Sigma}$ et $g(\cdot)$. Supposons que $(\mathbf{Y}_1, \mathbf{Y}_2)^\top \sim \mathcal{E}(g, 0, \mathbf{R})$, où \mathbf{R} est défini de sorte que

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix}.$$

Le Lemme 1 de Remillard (2012) nous indique que si \mathbf{R} est une matrice de corrélation de cette forme et que $C_{g,\mathbf{R}}$ est la copule associée à la distribution $\mathcal{E}(g, 0, \mathbf{R})$, alors la copule conditionnelle de \mathbf{Y}_2 sachant $\mathbf{Y}_1 = \mathbf{y}_1$ est donnée par $C_{g_2, \tilde{\boldsymbol{\Omega}}}$, où $\tilde{\boldsymbol{\Omega}}$ est la matrice

de corrélation $\boldsymbol{\Omega} = \mathbf{R}_{11} - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}$, et $g_2(\cdot)$ est donnée par

$$g_2(r) = \frac{g(r + \mathbf{y}_1^\top \mathbf{R}_{11}^{-1} \mathbf{y}_1)}{g_1(\mathbf{y}_1^\top \mathbf{R}_{11}^{-1} \mathbf{y}_1)} \quad \text{où} \quad g_1(r) = \int_{\mathbb{R}^d} g(\|\mathbf{y}_2\|^2 + r) d\mathbf{y}_2$$

De plus, notons $\mathbf{B} = \mathbf{R}_{21}\mathbf{R}_{11}^{-1}$, $|\mathbf{R}| = |\mathbf{R}_{11}||\boldsymbol{\Omega}|$ et

$\mathbf{y}^\top \mathbf{R}^{-1} \mathbf{y} = (\mathbf{y}_2 - \mathbf{B}\mathbf{y}_1)^\top \boldsymbol{\Omega}^{-1} (\mathbf{y}_2 - \mathbf{B}\mathbf{y}_1) + \mathbf{y}_1^\top \mathbf{R}_{11}^{-1} \mathbf{y}_1$. Alors, l'algorithme de simulation

pour générer une série chronologique U_1, \dots, U_n avec distribution stationnaire C_{g_1, \mathbf{R}_1} tel que $(U_{t-1}, U_t) \sim C_{g, \mathbf{R}}$ est le suivant.

Algorithme 1 : (Remillard, 2012)

1. Générer $\mathbf{X}_0 = (X_{01}, \dots, X_{0d}) \sim \mathcal{E}(g, 0, \mathbf{R}_1)$
2. Fixer $U_1 = (F_1(X_{01}), \dots, F_1(X_{0d}))$
3. Pour $t \in \{2, \dots, n\}$,
 - (a) Générer $V_t \sim \mathcal{E}(g_2, 0, \Omega)$
 - (b) Fixer $X_t = V_t + BX_{t-1}$ et $U_t = (F(X_{t1}), \dots, F(X_{td}))$

Processus markovien et copule archimédienne

Le deuxième algorithme concerne la classe des copules archimédiennes. Rappelons que la différence entre les copules archimédiennes se trouve dans le générateur de la copule, et que les copules de Gumbel, Clayton et Frank font partie de cette classe.

Soit C une copule archimédienne de dimension $(d_1 + d_2)$ avec générateur φ . Posons $A_u = C(\mathbf{u}, \mathbf{1})$, où $\mathbf{u} \in [0, 1]^{d_1}$. Alors, pour $k \in (0, 1]$, la fonction quantile associée à la probabilité que $V_j \leq k$ conditionnellement à $U = u$ est donnée par

$$Q_u(r) = \varphi^{-1}[h_{d_1}^{-1}(rh_{d_1}\{\varphi(A_u)\}) - \varphi(A_u)] \quad (3.2)$$

$$\text{où } h_k(r) = (-1)^k \frac{d^k}{dx^k} \varphi^{-1}(r).$$

Le Lemme 2 de Remillard (2012) énonce que quand $(U, V) \sim C_{d_1+d_2, \varphi}$, alors la copule associée à la distribution conditionnelle de V si $U = u \in (0, 1)^{d_1}$ est archimédienne avec un générateur défini pour $k \in (0, 1]$, par

$$\psi_u(k) = h_{d_1}^{-1}[kh_{d_1}\{\varphi(A_u)\} - \varphi(A_u)] \quad (3.3)$$

où $A_u = C_{d_1+d_2,\varphi}(\mathbf{u}, \mathbf{1})$.

Alors, pour générer une série chronologique de Markov, U_0, \dots, U_n avec distribution stationnaire $C_{d,\varphi}$ et une distribution conjointe $(U_{t-1}, U_t) \sim C_{2d,\varphi}$, l'algorithme ci-dessous peut être utilisé

Algorithme 2 : (Remillard, 2012)

1. Générer $U_0 \sim C_{d,\varphi}$
2. Pour $t \in \{1, \dots, n\}$
 - (a) Fixer $A_{U_{t-1}} = C_{2d,\varphi}(U_{t-1}, \mathbf{1})$
 - (b) Générer $V_t \sim C_{d,\psi_{U_{t-1}}}$, où ψ_u est défini en (3.3)
 - (c) Fixer $U_t = (U_{t,1}, \dots, U_{t,d})$, où $U_{t,k} = Q_{U_{t-1}}(V_{t,k})$ pour $k \in \{1, \dots, d\}$ et Q_u définie en (3.2).

Une fois la série produite à l'aide de l'Algorithme 1 pour les copules gaussienne et de Student, ou de l'Algorithme 2 pour les copules de Gumbel, Clayton et Frank, il ne reste plus qu'à appliquer une des méthodes d'estimation des paramètres vues à la Section 2.4. Il suffit d'appliquer l'algorithme N fois, afin d'obtenir N séries chronologiques de longueur p .

3.2 Copule bivariée à paramètre constant pour série chronologique univariée

Dans la section précédente, l'accent portait sur le fait que l'on peut obtenir une mesure de dépendance à travers plusieurs séries. Cette section représentait donc le cas où l'on possède trois séries pour des temps t distincts. Par exemple, le nombre de passagers

moyens d'un vol d'avion du Québec vers la France pour des compagnies A, B et C, pourrait représenter nos séries chronologiques étudiées. Trouver si la dépendance est forte entre les compagnies A, B et C pour un temps t quelconque devient alors un atout pour celles-ci. Effectivement, si seules les compagnies A et B possèdent une forte dépendance, alors ces deux compagnies ne devraient pas se soucier de la compagnie C pour définir leur prix de vols ou tout autre variables en lien avec le nombre de passagers moyens présent dans l'avion. Par contre, il ne faut pas pour autant oublier la compagnie C, même si la dépendance est plus faible.

Dans cette section, l'intérêt est porté sur la dépendance à l'intérieur d'une seule série chronologique. On s'intéressera donc à étudier une même variable à des temps différents. Par exemple, supposons que la compagnie aérienne A veut étudier la dépendance entre les nombres moyens de passagers sur ses vols du Québec vers la France aux temps $(t - 1)$ et t . Si cette dépendance est forte, alors cela signifie que ce qui s'est passé au temps $(t - 1)$ devrait se répéter au temps t . On obtient donc de l'information sur la tendance de la série temporelle. Cependant, il ne faut pas oublier qu'un effet aléatoire est présent dans une série. On ne peut donc pas totalement faire confiance aux valeurs obtenues pour la dépendance. L'analyse de la dépendance à l'intérieur d'une série chronologique présente donc plusieurs points communs avec l'analyse de la dépendance entre plusieurs séries.

Il serait possible d'étudier à l'intérieur d'une seule série chronologique s'il y a dépendance entre les temps $t, t - 1, t - 2, \dots$. Cependant, l'intérêt de ce mémoire porte sur des séries suivant un processus markovien d'ordre 1, d'où seules les dépendances entre deux temps successifs seront étudiées. D'ailleurs, la dépendance entre plusieurs séries chronologiques est strictement reliée à la dépendance présente dans chaque série individuellement. Pour X_t et Y_t deux séries chronologiques représentées par un

processus markovien d'ordre 1, notons

$$\mathbf{X}_t = \begin{pmatrix} X_{t-1} \\ X_t \end{pmatrix} \quad \text{et} \quad \mathbf{Y}_t = \begin{pmatrix} Y_{t-1} \\ Y_t \end{pmatrix}.$$

Ainsi, il est possible d'obtenir toutes les caractéristiques de ces deux séries temporelles à l'aide des couples (X_t, X_{t-1}) et (Y_t, Y_{t-1}) . D'ailleurs, la dépendance entre \mathbf{X}_t et \mathbf{Y}_t

peut être obtenue à l'aide de toutes les combinaisons que l'on peut former entre $(X_t, X_{t-1}, Y_t, Y_{t-1})$. Parmi les couples formés, on retrouve (X_t, Y_t) et (X_{t-1}, Y_{t-1}) qui nous informe sur la dépendance entre les deux séries chronologiques au même temps.

Deux autres couples sont représentés par (X_{t-1}, X_t) et (Y_{t-1}, Y_t) . Ces deux couples expliquent la dépendance à l'intérieur de chacune des séries. Il est donc fort intéressant d'étudier les séries chronologiques individuellement, afin d'obtenir toute l'information nécessaire pour l'analyse de la dépendance.

Soit une série chronologique (Y_1, Y_2, \dots, Y_n) à laquelle s'applique un processus markovien stationnaire d'ordre 1. Il suffit donc de déterminer la dépendance entre Y_t et Y_{t-1} . Dans le cas d'un modèle AR(1) de moyenne nulle, on peut identifier cette dépendance en estimant le paramètre a dans l'équation suivante :

$$Y_t = a \cdot Y_{t-1} + \epsilon_t$$

Sous les conditions de travail du processus AR(1), on peut montrer facilement que a représente la corrélation entre Y_t et Y_{t-1} :

$$\begin{aligned} \text{Corr}(Y_t, Y_{t-1}) &= \frac{\text{Cov}(Y_t, Y_{t-1})}{\sigma_{Y_t} \cdot \sigma_{Y_{t-1}}} \\ &= \frac{\text{Cov}(a \cdot Y_{t-1} + \epsilon_t, Y_{t-1})}{\sigma_{Y_t} \cdot \sigma_{Y_{t-1}}} \\ &= a \cdot \frac{\text{Cov}(Y_{t-1}, Y_{t-1})}{\sigma_{Y_t} \cdot \sigma_{Y_{t-1}}} \\ &= a \cdot \frac{\sigma^2}{\sigma^2} = a. \end{aligned}$$

Comme a est une constante, la dépendance entre (Y_1, Y_2) , (Y_2, Y_3) jusqu'à (Y_{n-1}, Y_n) est alors la même. Ainsi, estimer cette dépendance à l'aide d'une fonction copule et d'un paramètre θ constant résout aussi le problème. Estimer le paramètre a revient en fait à estimer le paramètre θ de la fonction copule à l'aide des méthodes vues précédemment.

Dans le cas d'une copule gaussienne (avec des lois normales comme distributions marginales), trouver le paramètre a équivaut à trouver le paramètre θ . En effet, pour une copule gaussienne $C(\cdot, \cdot; \theta)$ définie comme en (2.11), $\{\Phi^{-1}(F^*(Y_t))\}$ est un processus gaussien représenté par

$$\Phi^{-1}(F^*(Y_t)) = \theta \Phi^{-1}(F^*(Y_{t-1})) + \epsilon_t$$

où $\epsilon_t \sim N(0, 1 - \theta^2)$ et le bruit blanc au temps t est indépendant de Y_{t-1} . Si les distributions marginales sont des lois normales, on obtient que $F^*(\cdot) = \Phi(\cdot)$, ainsi que le modèle AR(1) suivant :

$$\Phi^{-1}(\Phi(Y_t)) = \theta \Phi^{-1}(\Phi(Y_{t-1})) + \epsilon_t$$

$$\text{d'où } Y_t = \theta \cdot Y_{t-1} + \epsilon_t.$$

Ce résultat nous indique clairement que le paramètre de la copule θ est égal au paramètre a pour ces conditions. Pour toutes lois autres que la normale, un processus de Markov d'ordre 1 peut être obtenu pour la copule gaussienne. Cependant, la simplicité des équations obtenues devient rapidement défunte.

Ainsi, pour tout temps t , toutes les propriétés statistiques attachées au modèle markovien d'ordre 1 proposé seront déterminées par la distribution conjointe de Y_t et Y_{t-1} . Les propriétés de la distribution conjointe peuvent être trouvées à l'aide des distributions marginales et de la fonction copule. De plus, les distributions marginales sont les mêmes aux temps t et $t - 1$. Le processus markovien d'ordre 1 peut donc être représenté par une copule avec la marginale F . L'avantage, comme dans le cas multivarié, est de pouvoir travailler sur la fonction copule et les marginales.

Une approche standard, comme pour le cas multivarié, est de travailler conditionnellement. Ainsi, si l'on connaît la densité conditionnelle de Y_t sachant Y_{t-1} , on peut estimer la distribution marginale de Y_{t-1} ou Y_t . Pour la densité conditionnelle

$h(\cdot|y_{t-1})$, on a l'égalité suivante :

$$h(y_t|y_{t-1}) = f(y_t) \cdot c(F(y_t), F(y_{t-1}); \theta)$$

où $f(\cdot)$ est la densité de la distribution marginale $F(\cdot)$ et $c(\cdot, \cdot; \theta)$ désigne la densité de copule avec paramètre θ . On remarque que comme annoncé, les distributions marginales sont les mêmes pour chaque temps de la série étudiée. De plus, la densité de cette distribution est par le fait même identique pour chaque temps. L'estimation pour la distribution marginale et du paramètre de la copule suit la même logique que pour la Section 3.1. On utilise donc les méthodes d'estimation vue à la Section 2.4, avec comme

restriction que $F(\cdot) = F_i(\cdot)$ pour $i = 1, 2, \dots, n$.

Jusqu'ici, l'emphase a été mise sur le fait que l'on peut utiliser la distribution conditionnelle afin d'estimer les paramètres. Par contre, il n'est pas nécessaire de calculer les résidus standardisés afin d'obtenir de bons estimateurs, comme pour l'approche décrite à la Section 3.1. Il est possible effectivement d'utiliser directement les méthodes d'estimation de la Section 2.4. L'ajustement d'un modèle paramétrique à la série chronologique n'étant pas nécessaire dans ces méthodes, calculer les résidus devient inutile. On obtient les estimateurs \hat{F} et $\hat{\theta}$ en appliquant, par exemple, la méthode IFM où à la première étape, F est estimé à l'aide de la distribution empirique redimensionnée et qu'à la deuxième étape, on estime le paramètre de la copule de la façon suivante :

$$(\text{Deuxième étape}) : \hat{\theta} = \arg \max \sum_{t=1}^n \log c(\hat{F}(y_t), \hat{F}(y_{t-1}); \theta).$$

3.3 Copule bivariée à paramètre dynamique pour série chronologique univariée

Jusqu'ici, le paramètre θ de la copule a toujours été considéré constant. Cependant, celui-ci pourrait être dépendant du temps dans certaines situations. De la même manière que la dépendance entre deux séries chronologiques pourrait augmenter ou diminuer avec le temps, la dépendance entre Y_t et Y_{t-1} pourrait être différente de celle entre Y_{t+100} et $Y_{t+100-1}$ pour une même série chronologique. Dans cette section, nous discuterons du changement d'un paramètre constant à un paramètre dépendant du temps, ou dynamique, pour la fonction copule.

La structure de dépendance peut être croissante ou décroissante à travers le temps. Dans un tel cas de non-stationnarité, l'estimation d'un paramètre de dépendance constant n'est plus une solution acceptable et l'équation du modèle AR(1) est remplacée par

$$Y_t = a_t \cdot Y_{t-1} + \epsilon_t.$$

Il s'en suit que pour la copule, il serait préférable d'estimer un paramètre θ_t dépendant du temps.

Estimation du paramètre dynamique

L'estimation d'un paramètre θ_t dépendant du temps peut s'effectuer à l'aide de méthodes similaires à celles utilisées dans le cas constant. Le choix de la fonction représentant θ_t n'est pas unique et chaque famille de copule possède des contraintes sur les paramètres. Par exemple, les copules gaussienne et de Student demandent que $\theta \in (-1, 1)$, la copule de Gumbel que $\theta \in [1, \infty)$, etc. Il en va de même pour les valeurs du paramètre θ_t choisi. Il serait intéressant de choisir ce paramètre de façon à pouvoir

comparer les différentes familles de copule entre elles même si les familles de copules n'ont pas les mêmes propriétés et les paramètres associés à celles-ci n'ont pas le même domaine de définition. Par contre, tel que mentionné au tableau 2.1, toutes les copules sont reliées entre elles par le tau de Kendall. L'approche à suivre est de représenter le tau de Kendall par une équation dépendante du temps, et d'y appliquer le tau de Kendall inverse pour obtenir un paramètre θ_t pour chaque copule. Dans ce mémoire, le tau de Kendall dépendant du temps, noté τ_t , est représenté par la fonction

$$\tau_t = \exp\left(\beta_0 + \beta_1 \cdot \frac{t}{T}\right) \quad (3.4)$$

où $t = 1, \dots, T$ avec T le temps du dernier élément de la série.

La fonction 3.4 choisie étant l'exponentielle d'une fonction linéaire, la fonction est croissante ou décroissante pour des paramètres β_0 et β_1 fixés. Cette fonction ne permet donc pas de traiter le cas où la dépendance augmente et puis diminue, ou le contraire. Il serait toutefois possible de traiter le cas où la dépendance est négative en ajoutant un signe négatif devant l'exponentielle de l'équation 3.4.

Notons que pour la copule de Frank, aucune forme explicite n'est disponible pour le tau de Kendall inverse. Par contre, dans le chapitre 4, une fonction du logiciel R sera utilisée afin de calculer le tau de Kendall inverse pour nos copules, incluant la copule de Frank. Il s'agit de la fonction BiCopTau2Par du package VineCopula.

L'estimation fournira $(\hat{\beta}_0, \hat{\beta}_1)$. La partie $\exp(\hat{\beta}_0)$ représente une estimée de la dépendance au début de la série ($t = 1$), alors que $\exp\left(\frac{\hat{\beta}_1}{T}\right)$ est l'incrément correspondant au changement dans la dépendance lors du passage du temps t au temps $t + 1$. Une estimée $\hat{\beta}_1 = 0$ signifie que la dépendance est constante dans le temps. Mais, qu'en est-il de $\hat{\beta}_1 = 0.5$ ou $\hat{\beta}_1 = -0.5$? Posséder l'écart-type pour nos paramètres estimés afin de construire un intervalle de confiance nous permettra de répondre à ces questions.

Copule	Tau de Kendall	Tau de Kendall inverse
Gaussienne	$\tau_t = \frac{2}{\pi} \arcsin \rho_t$	$\rho_t = \sin \left(\frac{\pi}{2} \cdot \tau_t \right)$
Student		
Gumbel	$\tau_t = 1 - \frac{1}{\theta_t}$	$\theta_t = \frac{1}{1 - \tau_t}$
Clayton	$\tau_t = \frac{\theta_t}{\theta_t + 2}$	$\theta_t = \frac{2\tau_t}{1 - \tau_t}$
Frank	$\tau_t = 1 + \frac{4}{\theta_t} [D_1(\theta_t) - 1]$	Pas de forme fermée

Tableau 3.1 – Tau de Kendall et inverse du tau de Kendall dynamique pour chacune des copules d'intérêt, avec $D_k(x) = \frac{k}{x} \int_0^x \frac{s^k}{e^s - 1} ds$

Dans la section qui suit, nous commencerons par montrer qu'asymptotiquement, les estimateurs des paramètres θ_t et F pour la copule sont convergents. Nous poursuivrons en détaillant un algorithme de simulation qui est basé sur ces résultats asymptotiques et nous permettra d'obtenir un intervalle de confiance pour nos paramètres d'intérêt.

Pour ce faire, nous adapterons la procédure de Nasri, Rémillard, Bouezmarni (2019)

[NRB19] à notre cas.

Convergence des estimateurs

Pour le paramètre d'une copule dépendant du temps, considérons une série non stationnaire pour laquelle les distributions marginales ne dépendent pas du temps. Ainsi, le paramètre de la copule peut changer avec le temps, mais les paramètres de la loi marginale restent les mêmes. Soient X_t et X_{t-1} avec $t \in \{1, \dots, T\}$ les variables d'intérêt. Les distributions marginales de X_t et X_{t-1} sont données par F et la copule à paramètre dépendant du temps est représentée par C_{θ_t} , avec c_{θ_t} comme densité de copule. La distribution conjointe est définie par H . Il est permis alors d'étudier une

seule série chronologique en posant des marginales équivalentes pour toutes les variables. Le théorème de Sklar nous indique, dans ce cas, que

$$H_t(x_t, x_{t-1}) = C_{\theta_t}(F(x_t), F(x_{t-1})).$$

Si F_n et $(\hat{\beta}_0, \hat{\beta}_1)$ sont des estimateurs convergents de F et θ_t , alors un estimateur convergent de $H_t(x_t, x_{t-1})$ est donné par

$$H_{n,t}(x_t, x_{t-1}) = C_{\theta_{n,t}}(F_n(x_t), F_n(x_{t-1})).$$

Pour obtenir les estimateurs convergents de F et θ_t , on peut utiliser la méthode semi-paramétrique présentée à la Section 2.4.

La distribution empirique redimensionnée fournit un estimateur des distributions marginales. Ainsi, on estime la distribution F par $F_n(x) = \frac{1}{n+1} \sum_{t=1}^n \mathbb{I}(X_t \leq x)$. De plus, il est connu que $\mathbb{F}_n(x) = n^{1/2} [F_n(x) - F(x)]$ converge en loi vers $\mathbb{B}[F(x)]$, où \mathbb{B} est un pont Brownien. Cette propriété est très intéressante parce qu'un pont brownien est un processus gaussien centré et continu, avec comme fonction de covariance

$$\text{Cov}(\mathbb{B}_j(m), \mathbb{B}_j(p)) = \min(m, p) - mp \text{ où } m, p \in [0, 1].$$

Pour estimer le paramètre de la copule, on peut appliquer la méthode de maximum de vraisemblance comme dans la méthode IFM. En posant $U_t = F(X_t)$ et $U_{t-1} = F(X_{t-1})$, et en utilisant les estimateurs $U_{n,t} = F_n(X_t)$, $U_{n,t-1} = F_n(X_{t-1})$ obtenus précédemment ainsi que $(\hat{\beta}_0, \hat{\beta}_1)$, on obtient l'estimateur suivant pour le paramètre de la copule

$$(\hat{\beta}_0, \hat{\beta}_1) = \arg \max \sum_{t=1}^n \log [c_{\theta_t}(U_{n,t}, U_{n,t-1})].$$

Il s'en suit que cet estimateur est \sqrt{n} -convergent si la densité de copule c_{θ_t} est suffisamment lisse, c'est-à-dire que sa dérivée première est bornée par une fonction intégrable. (Genest et al., 1995) [GGR95] [CL98]

En bout de courses, ces estimateurs nous permettent d'obtenir $H_{n,t}$. De plus, si les conditions du Théorème 1 de Nasri, Rémillard & Bouezmarni (2019) [NRB19] sont

respectées, on obtient la convergence de

$$\mathbb{H}_{n,t}(X_t, X_{t-1}) = n^{1/2} [H_{n,t}(X_t, X_{t-1}) - H_t(X_t, X_{t-1})] \text{ vers } \mathbb{H}_t(X_t, X_{t-1}).$$

Voyons comment adapter ce théorème au cas markovien d'ordre 1. Notons d'abord que l'on peut représenter le paramètre θ_t par une fonction $h_t\theta$. Par exemple, h_t pourrait être une

matrice de la forme $\begin{pmatrix} 1 & 0 \\ 0 & \frac{t}{T} \end{pmatrix}$. Ainsi, on peut noter que

$$\frac{\partial c_{\theta_t}(U_t, U_{t-1})}{\partial \theta_t} = h_t \dot{c}_{\theta_t}(U_t, U_{t-1})$$

et que $\frac{\partial^2 c_{\theta_t}(U_t, U_{t-1})}{\partial \theta_t^2} = h_t^2 \ddot{c}_{\theta_t}(U_t, U_{t-1})$

où \dot{c}_{θ_t} et \ddot{c}_{θ_t} représente le gradient et la matrice hessienne de la densité de copule par rapport à θ_t . Il est à noter que U_t correspond à u et que U_{t-1} correspond à v .

Notons aussi

$$\begin{aligned} \mathcal{W}_n &= n^{-1/2} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \\ \mathcal{J}_n &= n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1}) \dot{c}_{\theta_t}^\top(U_t, U_{t-1}) h_t}{c_{\theta_t}^2(U_t, U_{t-1})} \\ \mathcal{J}_n(x) &= \int_0^1 \left[n^{-1} \sum_{t=1}^n h_t^\top \left(\partial_v \dot{c}_{\theta_t}(u, F(x)) - \frac{\dot{c}_{\theta_t}(u, F(x)) \partial_v c_{\theta_t}(u, F(x))^\top}{c_{\theta_t}(u, F(x))} \right) f(x) \right] du \end{aligned}$$

Le théorème approprié à notre cas est présenté ci-dessous.

Théorème 3.1 *Supposons que la condition de Lindeberg est satisfaite pour la suite $\frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})}$, $t \in \{1, \dots, n\}$. Assumons aussi qu'uniformément sur \mathbb{R} , $(\mathcal{J}_n(x), \mathcal{J}_n)$ converge en probabilité vers $(\mathcal{J}(x), \mathcal{J})$, tel que \mathcal{J}^{-1} existe et que $n^{-1} \sum_{t=1}^n \frac{h_t^\top \ddot{c}_{\theta_t}(U_t, U_{t-1}) h_t}{c_{\theta_t}^2(U_t, U_{t-1})}$ converge en probabilité vers 0. Posons $\Theta_n = n^{1/2}(\theta_{n,t} - \theta_t)$. Alors, quand $n \rightarrow \infty$, on a que $(\mathcal{W}_n, \Theta_n)$ converge en loi vers (\mathcal{W}, Θ) , où $\mathcal{W} \sim N(0, \mathcal{J})$ et $\Theta = \mathcal{J}^{-1}(\mathcal{W} + \int_{\mathbb{R}} \mathcal{J}(x) dx)$.*

Preuve :

Le but étant de maximiser

$$\log c_{\theta_t}(U_{n,t}, U_{n,t-1})$$

par rapport à θ , on cherche donc la valeur de θ qui satisfait

$$0 = n^{-1/2} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_{n,t}, U_{n,t-1})}{c_{\theta_t}(U_{n,t}, U_{n,t-1})} \quad (3.5)$$

En appliquant la méthode delta (J.L.Doob 1935) et (Robert Dorfman 1938), on obtient

$$\begin{aligned} 0 &= n^{-1/2} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} + \frac{\partial}{\partial \theta} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (\theta_{n,t} - \theta_t) \\ &+ \frac{\partial}{\partial U_t} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (U_{n,t} - U_t) \\ &+ \frac{\partial}{\partial U_{t-1}} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (U_{n,t-1} - U_{t-1}) \\ &= \mathcal{W}_n + I_{1,n} + I_{2,n} + I_{3,n}. \end{aligned}$$

Le premier terme de cette équation est

$$\mathcal{W}_n = n^{-1/2} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})}.$$

Le deuxième terme, $I_{2,n}$, nous donne

$$\begin{aligned}
I_{2,n} &= \frac{\partial}{\partial \theta} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (\theta_{n,t} - \theta_t) \\
&= n^{-1} \sqrt{n} (\theta_{n,t} - \theta_t) \sum_{t=1}^n \frac{\partial}{\partial \theta} \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \\
&= n^{-1} \sqrt{n} (\theta_{n,t} - \theta_t) \sum_{t=1}^n \frac{\partial}{\partial \theta} \left[h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1}) \cdot \frac{1}{c_{\theta_t}(U_t, U_{t-1})} \right] \\
&= n^{-1} \sqrt{n} (\theta_{n,t} - \theta_t) \sum_{t=1}^n h_t^\top \left[\frac{h_t \ddot{c}_{\theta_t}(U_t, U_{t-1}) c_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})} \right. \\
&\quad \left. - \frac{\dot{c}_{\theta_t}(U_t, U_{t-1}) h_t \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})} \right] \\
&= -n^{-1} \sqrt{n} (\theta_{n,t} - \theta_t) \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1}) \dot{c}_{\theta_t}^\top(U_t, U_{t-1}) h_t}{c_{\theta_t}^2(U_t, U_{t-1})} + o_p(1) \quad (*) \\
&= -n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1}) \dot{c}_{\theta_t}^\top(U_t, U_{t-1}) h_t}{c_{\theta_t}^2(U_t, U_{t-1})} \sqrt{n} (\theta_{n,t} - \theta_t) + o_p(1)
\end{aligned}$$

(*) Le Théorème 3.1 suppose que $n^{-1} \sum_{t=1}^n \frac{h_t^2 \ddot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})}$ converge en probabilité vers 0.

En ce qui concerne le troisième terme, $I_{3,n}$,

$$\begin{aligned}
I_{3,n} &= \frac{\partial}{\partial U_t} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} \right] \sqrt{n} (U_{n,t} - U_t) \\
&= n^{-1} \sqrt{n} (U_{n,t} - U_t) \sum_{t=1}^n h_t^\top \frac{\partial}{\partial U_t} \left[\frac{\dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} \right] \\
&= n^{-1} \sqrt{n} (U_{n,t} - U_t) \sum_{t=1}^n h_t^\top \left[\frac{\partial}{\partial U_t} [\dot{c}_\theta(U_t, U_{t-1})] \cdot \frac{1}{c_\theta(U_t, U_{t-1})} \right. \\
&\quad \left. + \dot{c}_\theta(U_t, U_{t-1}) \cdot \frac{\partial}{\partial U_t} \left[\frac{1}{c_\theta(U_t, U_{t-1})} \right] \right] \\
&= n^{-1} \sqrt{n} (U_{n,t} - U_t) \sum_{t=1}^n h_t^\top \left[\left[\frac{\frac{\partial}{\partial U_t} \dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} \right] \right. \\
&\quad \left. - \dot{c}_\theta(U_t, U_{t-1}) \cdot \frac{1}{c_\theta^2(U_t, U_{t-1})} \cdot \frac{\partial}{\partial U_t} c_\theta(U_t, U_{t-1}) \right] \\
&= n^{-1} \sqrt{n} (U_{n,t} - U_t) \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_t} \dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} - \frac{\dot{c}_\theta(U_t, U_{t-1}) \frac{\partial}{\partial U_t} c_\theta(U_t, U_{t-1})}{c_\theta^2(U_t, U_{t-1})} \right] \\
&= n^{-1} \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_t} \dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} - \frac{\dot{c}_\theta(U_t, U_{t-1}) \frac{\partial}{\partial U_t} c_\theta(U_t, U_{t-1})}{c_\theta^2(U_t, U_{t-1})} \right] \sqrt{n} (U_{n,t} - U_t).
\end{aligned}$$

Pour le quatrième terme, la dérivée se fait par rapport à U_{t-1} au lieu de U_t , mais ces deux variables sont représentées par la même distribution marginale. On obtient donc

en procédant de manière similaire

$$\begin{aligned}
&\frac{\partial}{\partial U_{t-1}} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} \right] \sqrt{n} (U_{n,t-1} - U_{t-1}) \\
&= n^{-1} \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_{t-1}} \dot{c}_\theta(U_t, U_{t-1})}{c_\theta(U_t, U_{t-1})} - \frac{\dot{c}_\theta(U_t, U_{t-1}) \frac{\partial}{\partial U_{t-1}} c_\theta(U_t, U_{t-1})}{c_\theta^2(U_t, U_{t-1})} \right] \sqrt{n} (U_{n,t-1} - U_{t-1}).
\end{aligned}$$

En regroupant les quatre termes à nouveau, il s'en suit que

$$\begin{aligned}
0 &= n^{-1/2} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} + \frac{\partial}{\partial \theta} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (\theta_{n,t} - \theta_t) \\
&+ \frac{\partial}{\partial U_t} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (F_n - F) \\
&+ \frac{\partial}{\partial U_{t-1}} \left[n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} \right] \sqrt{n} (F_n - F) \\
&= \mathcal{W}_n - n^{-1} \sum_{t=1}^n \frac{h_t^\top \dot{c}_{\theta_t}(U_t, U_{t-1}) \dot{c}_{\theta_t}^\top(U_t, U_{t-1}) h_t}{c_{\theta_t}^2(U_t, U_{t-1})} \sqrt{n} (\theta_{n,t} - \theta_t) \\
&+ n^{-1} \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_t} \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} - \frac{\dot{c}_{\theta_t}(U_t, U_{t-1}) \frac{\partial}{\partial U_t} c_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})} \right] \sqrt{n} (F_n - F) \\
&+ n^{-1} \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_{t-1}} \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} - \frac{\dot{c}_{\theta_t}(U_t, U_{t-1}) \frac{\partial}{\partial U_{t-1}} c_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})} \right] \sqrt{n} (F_n - F) \\
&+ o_p(1) \\
&= \mathcal{W}_n - \mathcal{J}_n \Theta_n + n^{-1} \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_t} \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} - \frac{\dot{c}_{\theta_t}(U_t, U_{t-1}) \frac{\partial}{\partial U_t} c_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})} \right] \mathbb{F}_n \\
&+ n^{-1} \sum_{t=1}^n h_t^\top \left[\frac{\frac{\partial}{\partial U_{t-1}} \dot{c}_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}(U_t, U_{t-1})} - \frac{\dot{c}_{\theta_t}(U_t, U_{t-1}) \frac{\partial}{\partial U_{t-1}} c_{\theta_t}(U_t, U_{t-1})}{c_{\theta_t}^2(U_t, U_{t-1})} \right] \mathbb{F}_n + o_p(1) \\
&= \mathcal{W}_n - \mathcal{J} \Theta_n \\
&+ \int_0^1 \int_{\mathbb{R}} \left[n^{-1} \sum_{t=1}^n h_t^\top \left(\partial_v \dot{c}_{\theta_t}(u, F(x_{t-1})) - \frac{\dot{c}_{\theta_t}(u, F(x)) \partial_v c_{\theta_t}(u, F(x))}{c_{\theta_t}(u, F(x))} \right) f(x) \right] du dx \\
&+ \left[n^{-1} \sum_{t=1}^n h_t^\top \int_{(0,1)^2} \left(\partial_u \dot{c}_{\theta_t}(u, v) - \frac{\dot{c}_{\theta_t}(u, v) \partial_u c_{\theta_t}(u, v)}{c_{\theta_t}(u, v)} \right) F \circ F^{-1}(u) du dv \right] + o_p(1) \\
&= \mathcal{W}_n + \int_{\mathbb{R}} \mathcal{J}_n(x) dx - \mathcal{J} \Theta_n + o_p(1).
\end{aligned}$$

Finalement,

$$\mathcal{J}\Theta_n = \mathcal{W}_n + \int_{\mathbb{R}} \mathcal{J}_n(x)dx + o_p(1)$$

et

$$\Theta_n = \mathcal{J}^{-1} \left(\mathcal{W}_n + \int_{\mathbb{R}} \mathcal{J}_n(x)dx + o_p(1) \right).$$

En conclusion, $(\mathcal{W}_n, \Theta_n)$ converge en loi vers (\mathcal{W}, Θ) .

Si les conditions du Théorème 3.1 sont respectées, le processus $\mathbb{H}_{n,t}(X_t, X_{t-1}) = n^{1/2} [H_{n,t}(X_t, X_{t-1}) - H_t(X_t, X_{t-1})]$ converge en loi, lorsque $n \rightarrow \infty$, vers un processus gaussien centré et continu \mathbb{H}_t , tel que

$$\begin{aligned} \mathbb{H}_t(x_t, x_{t-1}) = & \Theta^\top h_t^\top \dot{C}_{\theta_t}(F(x_t), F(x_{t-1})) + F(x_t) \partial_u C_{\theta_t}(F(x_t), F(x_{t-1})) \\ & + F(x_{t-1})^\top \nabla_v C_{\theta_t}(F(x_t), F(x_{t-1})). \end{aligned}$$

L'estimateur pour la distribution conjointe $H_{n,t}(x_t, x_{t-1}) = C_{\theta_{n,t}}(F_n(X_t), F_n(X_{t-1}))$ converge donc asymptotiquement, lorsque $n \rightarrow \infty$, vers la vraie distribution

$$H(x_t, x_{t-1}).$$

Dans la prochaine section, nous construirons un algorithme afin de simuler des séries chronologiques, dans le but de calculer un intervalle de confiance pour nos paramètres d'intérêt. L'algorithme servira à vérifier nos résultats précédents à propos de la convergence asymptotique et des résultats de simulation seront présentés au chapitre 4.

3.3.1 Algorithme de simulation

L'algorithme décrit dans cette section comporte deux parties. La première présente la simulation de N séries chronologiques dont la dépendance (l'autocorrélation) est

fonction du paramètre $\theta_{t,0}$. Dans la deuxième partie, on propose un algorithme pour construire un intervalle de confiance pour le paramètre $\theta_{t,0}$ de la copule.

Algorithme 1 :

1. Pour $i = 1, \dots, N$

- (a) Générer $(T + 20)$ variables uniformes indépendantes $u_1^{(i)}, \dots, u_{T+20}^{(i)}$.
- (b) Posons $x_1^{(i)} = u_1^{(i)}$ et $k_u(v) = \frac{\partial C(u, v; \theta_{t,0})}{\partial u}$ avec $\theta_{t,0}$, le paramètre supposé de la copule.
- (c) Pour $j = 2, \dots, (T + 20)$, posons $x_j^{(i)} = k_{x_{j-1}^{(i)}}^{-1}(u_j^{(i)})$.
- (d) Enlever les 10 premières et les 10 dernières données.

On obtient N séries de T observations $(x_1^{(1)}, \dots, x_T^{(1)}), \dots, (x_1^{(N)}, \dots, x_T^{(N)})$.

Cet algorithme fonctionne pour plusieurs structures de dépendance, dont les cinq familles de copules étudiées dans ce mémoire. De plus, pour enlever tout effet de biais des premières et des dernières valeurs des séries simulées, les 10 premières et 10 dernières valeurs de chaque série ont été enlevées.

Soient $(x_1^{(1)}, \dots, x_T^{(1)}), \dots, (x_1^{(N)}, \dots, x_T^{(N)})$, nos N séries chronologiques de T observations obtenues à l'aide de l'Algorithme 1.

Algorithme 2 :

1. Pour $i = 1, \dots, N$,

calculer $\hat{\theta}_{t,0}^{(i)} = \arg \max \sum_{t=2}^T \log c(F_n(x_t^{(i)}), F_n(x_{t-1}^{(i)}); \theta_{t,0})$.

2. Calculer $\mu(\hat{\theta}_{t,0}) = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_{t,0}^{(i)}$

3. Calculer $\sigma(\hat{\theta}_{t,0}) = \sqrt{\sum_{i=1}^N \frac{[\hat{\theta}_{t,0}^{(i)} - \mu(\hat{\theta}_{t,0})]^2}{N - 1}}$

4. Obtenir l'intervalle de confiance au pourcentage $100(1 - \alpha)$, avec $\alpha \in (0, 1)$.

Par exemple, en supposant que la distribution des données est normale, on peut obtenir l'intervalle de confiance à $100(1 - \alpha)\%$ suivant : $\mu(\hat{\theta}_{t,0}) \pm z_{\alpha/2} \cdot \sigma(\hat{\theta}_{t,0})$. On peut aussi obtenir un intervalle de confiance grâce aux quantiles empiriques. De plus, même si les résultats de cette section seront applicables, notons encore que l'algorithme utilise une distribution marginale semblable pour x_t et x_{t-1} , et que cette distribution ne varie pas avec le temps.

CHAPITRE 4

Applications

L'étude de la dépendance à travers une série chronologique, non supposée constante, est au coeur de ce mémoire. Pour évaluer si l'estimation du paramètre de la copule dépendant du temps donne de bons résultats, nous considérerons deux exemples. D'abord, nous analyserons un cas théorique, où le nombre d'observations dans la série chronologique est variant selon les cas proposés. Des estimations pour des séries avec différentes tailles d'observations seront présentées. Par la suite, nous considérerons des séries de prix du blé pour cinq pays, provenant du site officiel de *Food and Agriculture Organization (FAO)*. Nous examinerons si la structure de dépendance de ces séries est significativement dépendante du temps, ou si un paramètre à valeur constante est suffisant pour représenter le paramètre de la dépendance, θ . Notons que les estimateurs obtenus sont analysés à l'aide du biais par rapport aux vrais paramètres, ainsi qu'à l'écart-type. Un estimateur est de bonne qualité s'il possède un biais minimal et un plus petit écart-type. De plus, pour les tableaux qui suivent, les titres de colonnes « I.C. (β) (N) » et « I.C. (β) (Q) » signifie respectivement qu'il s'agit de l'intervalle de confiance obtenu de la loi normale et de l'intervalle de confiance calculé à partir des quantiles

empiriques pour le paramètre β .

4.1 Applications théoriques

Le but de cette section est de travailler sur le paramètre θ_t . L'intérêt est porté sur la croissance ou la décroissance de la dépendance. L'équation (3.4) sera à nouveau utilisée pour représenter les valeurs de τ_t .

Pour obtenir les valeurs de θ_t , il suffit d'utiliser le tau de Kendall inverse pour chacune de nos copules, en remplaçant τ_t par son équivalence dans l'équation (3.4). Rappelons que la fonction `BiCopTau2Par` du package `CopulaVine` est utilisée afin de calculer le tau de Kendall inverse dans le logiciel R. Ici, l'emphasis est mise sur le paramètre β_1 afin de savoir à partir de quelle valeur la variation de la dépendance par rapport au temps est significative, c'est-à-dire que l'intervalle de confiance est strictement positif ou négatif. Nous avons dès lors appliqué l'**Algorithme 1** avec $N = 500$,

$$n = \{120, 220, 520\} \text{ et } \beta_0 = -1.$$

De cette façon, 500 séries de 500, 200 et 100 observations seront obtenues pour les trois cas étudiés. Pour chaque série, la dépendance initiale est de $\tau_t = \exp(-1) = 0.3679$. On

cherche à partir de quelle valeur de β_1 on obtient une dépendance dynamique significative dans chaque cas. La copule étudiée ici est la gaussienne. L'**Algorithme 2** est appliqué afin de créer un intervalle de confiance à partir des 500 séries. Dans les tableaux qui suivront, la distribution marginale F utilisée est une distribution normale.

L'estimation de F est non-paramétrique et F_n est faite à partir de la distribution empirique redimensionnée. Tous les intervalles ont été construits avec un niveau de confiance de 95%. Les résultats obtenus sont présentés au tableau 4.1.

Comme il y a une augmentation de la dépendance au cours du temps, les valeurs testées

β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
0.1	-1.0029	0.0999	[-1.2525,-0.7534]	[-0.2767,0.4765]	[-1.2661,-0.7717]	[-0.2736,0.4774]
0.2	-0.9978	0.1937	[-1.2364,-0.7593]	[-0.1462,0.5336]	[-1.2535,-0.7921]	[-0.1264,0.5642]
0.3	-0.9958	0.2900	[-1.2226,-0.7690]	[-0.0138,0.5939]	[-1.2436,-0.7848]	[-0.0144,0.5925]
0.31	-1.0010	0.3026	[-1.2269,-0.7751]	[-0.0065,0.6117]	[-1.2425,-0.7813]	[0.0174,0.6118]
0.32	-1.0053	0.3314	[-1.2418,-0.7688]	[0.0141,0.6487]	[-1.2617,-0.7866]	[0.0010,0.6666]
0.33	-1.0045	0.3286	[-1.2407,-0.7684]	[0.0170,0.6403]	[-1.2582,-0.7886]	[0.0291,0.6526]
0.34	-1.0049	0.3368	[-1.2298,-0.7801]	[0.0372,0.6364]	[-1.2532,-0.7975]	[0.0594,0.6423]
0.35	-0.9945	0.3490	[-1.2151,-0.7740]	[0.0626,0.6354]	[-1.2348,-0.7816]	[0.0764,0.6263]
0.4	-1.0014	0.4046	[-1.2219,-0.7808]	[0.1177,0.6916]	[-1.2417,-0.8029]	[0.1368,0.6822]
0.5	-1.0045	0.5056	[-1.2215,-0.7875]	[0.2494,0.7618]	[-1.2315,-0.8109]	[0.2785,0.7683]
0.6	-0.9936	0.5924	[-1.1983,-0.7889]	[0.3618,0.8229]	[-1.2188,-0.8070]	[0.3747,0.8441]
0.7	-0.9886	0.6913	[-1.1949,-0.7823]	[0.4740,0.9087]	[-1.2130,-0.7857]	[0.4803,0.9257]
0.8	-1.0075	0.8089	[-1.2331,-0.7819]	[0.5781,1.0398]	[-1.2425,-0.7939]	[0.5926,1.0666]
0.9	-1.0092	0.9028	[-1.2477,-0.7707]	[0.6541,1.1516]	[-1.2369,-0.7648]	[0.6522,1.1494]

Tableau 4.1 – Intervalles de confiance (I.C.) pour les paramètres β_0 et β_1 dans le cas de la copule gaussienne avec $\beta_0 = -1$ fixe et $N = 500$ séries de $T = 500$ observations

pour β_1 sont positives. Pour chacun des deux paramètres, un intervalle de confiance utilisant la loi normale ainsi qu'un intervalle de confiance utilisant les quantiles empiriques ont été construits. La marge d'erreur des intervalles de confiance du second type (quantiles) semble généralement plus petite. Ensuite, malgré la modification de la valeur du paramètre β_1 , les intervalles de confiance pour β_0 changent à peine. Un changement au paramètre β_1 n'a donc pas d'effet sur l'estimation du paramètre β_0 . Pour $\beta_1 = 0.31$, seul l'intervalle de confiance calculé à partir des quantiles pour β_1 est strictement positif. Ce n'est qu'à partir de $\beta_1 = 0.32$ que les intervalles des deux types sont strictement positifs. Cependant, il faut garder en tête que la valeur de 0.32 obtenue pour β_1 n'est valide que si $\beta_0 = -1$, $N = 500$ et $T = 500$.

Les résultats présentés au tableau 4.2 où $T = 200$ vont nous permettre d'étudier l'effet de la longueur de séries sur la significativité de la dépendance du temps étudiée. Comme au tableau 4.1, les intervalles de confiance utilisant les quantiles empiriques

β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
0.3	-1.0142	0.3123	[-1.3783,-0.6501]	[-0.1825,0.8070]	[-1.4155,-0.7175]	[-0.1454,0.7836]
0.4	-1.0061	0.4043	[-1.3702,-0.6421]	[-0.0591,0.8678]	[-1.4227,-0.7006]	[-0.0314,0.8937]
0.41	-1.0024	0.4135	[-1.3547,-0.6501]	[-0.0512,0.8783]	[-1.4028,-0.7165]	[-0.0397,0.9250]
0.42	-1.0020	0.4217	[-1.3584,-0.6455]	[-0.0419,0.8854]	[-1.3759,-0.7039]	[-0.0033,0.8984]
0.43	-0.9825	0.4137	[-1.3252,-0.6399]	[-0.0338,0.8613]	[-1.3891,-0.6718]	[-0.0130,0.8966]
0.44	-0.9989	0.4388	[-1.3422,-0.6556]	[-0.0223,0.8998]	[-1.3888,-0.6802]	[-0.0496,0.9116]
0.45	-1.0039	0.4550	[-1.3628,-0.6449]	[0.0163,0.8938]	[-1.3964,-0.6887]	[0.0164,0.9491]
0.46	-1.0041	0.4607	[-1.3637,-0.6445]	[0.0020,0.9194]	[-1.4124,-0.6659]	[0.0211,0.9269]
0.5	-0.9997	0.4978	[-1.3593,-0.6401]	[0.0619,0.9336]	[-1.3964,-0.6934]	[0.0852,0.9429]

Tableau 4.2 – Intervalles de confiance (I.C.) pour les paramètres β_0 et β_1 dans le cas de la copule gaussienne avec $\beta_0 = -1$ fixe et $N = 500$ séries de $T = 200$ observations

versus les qualités basées sur la loi normale pour β_0 ne semble pas changer en fonction de β_1 . De plus, l'intervalle de confiance utilisant les quantiles empiriques semble répondre à nouveau sensiblement plus rapidement à la positivité de β_1 . Cependant, ce n'est qu'à partir de $\beta_1 = 0.45$ que l'on obtient la significativité pour les deux intervalles de confiance. Ce résultat est tout à fait logique, étant donné qu'une diminution du nombre d'observations par série conduit à des estimateurs moins performants (à la dispersion plus grande) pour chaque série. Ainsi, l'écart-type des estimations devient plus grand et un β_1 plus grand est nécessaire pour obtenir un intervalle de confiance strictement positif. Le tableau 4.3 où $T = 100$ confirme que lorsque T diminue, la valeur de β_1 pour détecter une dépendance variante à travers le temps augmente.

Les remarques faites pour les tableaux 4.1 et 4.2 s'appliquent à la différence entre les intervalles de confiance. Par contre, les intervalles de confiance pour β_0 ne sont plus aussi stables que lorsque T était plus élevé. Cela peut être dû au changement dans la valeur du paramètre β_1 , mais aussi au fait que lorsque T est petit, la convergence des estimateurs n'est plus garantie de la même manière. Avec $T = 100$, on observe qu'à partir de $\beta_1 = 0.59$, on obtient des intervalles de confiance strictement positifs.

β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
0.3	-1.0255	0.3227	[-1.5772,-0.4739]	[-0.4487,1.0941]	[-1.7040,-0.5839]	[-0.4236,1.1799]
0.4	-0.9751	0.3677	[-1.4868,-0.4633]	[-0.3103,1.0457]	[-1.5118,-0.5671]	[-0.2839,1.0338]
0.5	-0.9957	0.5068	[-1.5000,-0.4914]	[-0.1266,1.1403]	[-1.5965,-0.5725]	[-0.0680,1.1810]
0.55	-0.9883	0.5393	[-1.5046,-0.4720]	[-0.0805,1.1592]	[-1.6059,-0.5753]	[-0.0025,1.2905]
0.56	-0.9825	0.5496	[-1.4779,-0.4871]	[-0.0420,1.1413]	[-1.6024,-0.5978]	[0.00515,1.2317]
0.57	-0.9928	0.5714	[-1.5093,-0.4762]	[-0.0149,1.1578]	[-1.5737,-0.5839]	[0.0649,1.1746]
0.58	-0.9871	0.5690	[-1.4698,-0.5045]	[-0.0018,1.1398]	[-1.5298,-0.5810]	[0.0505,1.1915]
0.59	-1.0035	0.6050	[-1.5252,-0.4818]	[0.0048,1.2051]	[-1.6170,-0.5751]	[0.0581,1.2336]
0.6	-0.9917	0.5961	[-1.4825,-0.5009]	[0.0183,1.1738]	[-1.6030,-0.5896]	[0.0548,1.2418]

Tableau 4.3 – Intervalles de confiance (I.C.) pour les paramètres β_0 et β_1 dans le cas de la copule gaussienne avec $\beta_0 = -1$ fixe et $N = 500$ séries de $T = 100$ observations

Jusqu'ici, seule la copule gaussienne a été considérée. Or, tel que mentionné à la Section 2.4 à propos de l'estimation des paramètres de la copule, maximiser une fonction de vraisemblance en utilisant un seul type de copule n'est pas suffisant. Effectivement, chaque famille de copules offre des estimateurs $\hat{\beta}_0$ et $\hat{\beta}_1$ différents en appliquant une même méthode d'estimation. Le problème revient donc à trouver la copule qui offre la meilleure estimation pour une série chronologique donnée. Afin de comparer ces différentes copules, les critères du AIC ou du BIC peuvent être appliqués. Cependant, avant même de tester quelle copule est la meilleure dans un cas donné, il peut être intéressant de comparer l'estimation de la fonction τ_t pour ces différentes familles en utilisant des paramètres β_0 et β_1 semblables.

Au tableau 4.4 sont rapportés des résultats d'estimation pour cinq familles de copules.

Deux estimations par copule ont été effectuées, avec comme valeurs des paramètres $\beta_0 = -0.5$ et $\beta_1 = 0.2$. Le paramètre β_1 étant positif, la dépendance dans le temps est croissante. Le nombre de séries simulées afin d'obtenir les estimateurs est $N = 500$, avec $T = 500$ observations par série. Rappelons que pour des valeurs égales de β_0 et β_1 , la dépendance à travers le temps est la même pour les cinq copules.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
Gauss	-0.5129	0.1987	[-0.6298,-0.3961]	[0.0712,0.3263]	[-0.6321,-0.4004]	[0.0814,0.3237]
	-0.5150	0.1993	[-0.6414,-0.3886]	[0.0703,0.3283]	[-0.6423,-0.3979]	[0.0768,0.3376]
Student	-0.5160	0.1982	[-0.6600,-0.3720]	[0.0471,0.3494]	[-0.6669,-0.3876]	[0.0571,0.3677]
	-0.5287	0.2101	[-0.6809,-0.3766]	[0.0517,0.3685]	[-0.7038,-0.3978]	[0.0594,0.3790]
Gumbel	-0.5428	0.2204	[-0.7295,-0.3560]	[0.0387,0.4021]	[-0.7179,-0.3623]	[0.0385,0.4231]
	-0.5317	0.2072	[-0.7205,-0.3428]	[0.0208,0.3936]	[-0.7439,-0.3600]	[0.0163,0.4155]
Clayton	-0.5909	0.2309	[-0.9167,-0.2652]	[-0.0867,0.5486]	[-0.9260,-0.2695]	[-0.0920,0.5430]
	-0.5895	0.2188	[-0.9079,-0.2710]	[-0.0784,0.5159]	[-0.9233,-0.2818]	[-0.1095,0.5003]
Frank	-0.5096	0.1926	[-0.6114,-0.4077]	[0.0479,0.3373]	[-0.6197,-0.4129]	[0.0427,0.3262]
	-0.5119	0.1991	[-0.6188,-0.4049]	[0.0573,0.3409]	[-0.6271,-0.4136]	[0.0592,0.3538]

Tableau 4.4 – Estimation des paramètres pour différentes familles de copule obtenue à l’aide de $N = 500$ séries de $T = 500$ observations, pour $\beta_0 = -0.5$ et $\beta_1 = 0.2$

Les copules gaussienne, de Student et de Frank semblent être les mieux estimées au tableau 4.4. Les résultats les moins bons sont obtenus pour la copule de Clayton, suivi de la copule de Gumbel. La méthode utilisée (IFM) semble donner des estimés pour β_0 toujours inférieurs à la vraie valeur. Cependant, comme on s’intéresse à modéliser la dépendance dynamique dans le temps à l’aide de copules, β_1 est le paramètre d’intérêt et semble bien être estimé pour toutes les copules, bien que l’estimation pour la copule de Clayton et Gumbel offre des résultats moins encourageants. La longueur des intervalles de confiance nous indiquent si la variance de l’estimateur correspondant est élevée ou non. En comparant les intervalles de confiance de β_1 pour chaque copule, on remarque que la copule gaussienne fournit les estimations les plus précises, suivie par la copule de Frank, puis celle de Student. Les copules de Gumbel et celle de Clayton ferment la marche. Pour toutes les copules sauf celle de Clayton (dont le manque de précision dans l’estimation la place dans une classe à part), obtenir des valeurs pour les estimateurs près des valeurs des paramètres $\beta_0 = -0.5$ et $\beta_1 = 0.2$ signifie que la dépendance est significativement dépendante du temps. Effectivement, les intervalles de confiance pour chacune de ces copules sauf la Clayton sont strictement positifs.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
Gauss	-0.3220	-0.1957	[-0.4034,-0.2406]	[-0.3239,-0.0675]	[-0.4075,-0.2449]	[-0.3212,-0.0617]
	-0.3209	-0.2029	[-0.4128,-0.2406]	[-0.3320,-0.0829]	[-0.4061,-0.2357]	[-0.3338,-0.0719]
Student	-0.3297	-0.2047	[-0.4289,-0.2305]	[-0.3691,-0.0403]	[-0.4393,-0.2402]	[-0.3718,-0.0557]
	-0.3280	-0.2084	[-0.4227,-0.2332]	[-0.3733,-0.0435]	[-0.4203,-0.2309]	[-0.3963,-0.0474]
Gumbel	-0.3347	-0.2102	[-0.4525,-0.2169]	[-0.3969,-0.0234]	[-0.4637,-0.2222]	[-0.4284,-0.0462]
	-0.3377	-0.2105	[-0.4522,-0.2232]	[-0.3887,-0.0323]	[-0.4621,-0.2256]	[-0.3901,-0.0345]
Clayton	-0.3718	-0.2423	[-0.5574,-0.1862]	[-0.5553,0.0706]	[-0.5760,-0.2011]	[-0.5536,0.0871]
	-0.3763	-0.2217	[-0.5635,-0.1890]	[-0.5302,0.0868]	[-0.5952,-0.2103]	[-0.5091,0.1197]
Frank	-0.3270	-0.1935	[-0.4106,-0.2434]	[-0.3390,-0.0481]	[-0.4176,-0.2488]	[-0.3437,-0.0550]
	-0.3279	-0.1895	[-0.4103,-0.2456]	[-0.3423,-0.0367]	[-0.4086,-0.2543]	[-0.3399,-0.0343]

Tableau 4.5 – Estimation des paramètres pour différentes familles de copule obtenue à l’aide de $N = 500$ séries de $T = 500$ observations, pour $\beta_0 = -0.3$ et $\beta_1 = -0.2$

On observe au tableau 4.5 un tableau construit de façon identique. Cependant, des paramètres $\beta_0 = -0.3$ et $\beta_1 = -0.2$ ont été utilisés pour ces simulations, afin d’observer une dépendance décroissante à travers le temps. Voyons voir si les résultats semblent les mêmes lorsque la dépendance est décroissante à travers le temps. Tout comme pour le tableau 4.4, on remarque que les valeurs de β_0 estimées semblent être inférieures à la valeur réelle. Les estimateurs de β_1 sont adéquats, malgré que tout comme lorsque $\beta_1 > 0$, la copule de Clayton possède les valeurs les plus éloignées. Les intervalles de confiance nous donnent sensiblement le même ordre pour la variance des estimations de nos copules que pour le tableau 4.4. Les intervalles de confiance pour β_1 étant toutes négatives (mis à part la copule de Clayton), on peut donc affirmer que la dépendance à travers le temps devrait changer, voir même diminuer pour être plus précis, étant donné que $\beta_1 < 0$.

Une des raisons principales des résultats obtenus pour la copule de Clayton, ainsi que la copule de Gumbel est cachée dans le paramètre de la copule directement. Le tau de Kendall inverse nous permet de définir le paramètre de la copule en fonction du tau de Kendall, et ainsi, si la valeur donnée au tau de Kendall ne respecte pas son domaine de

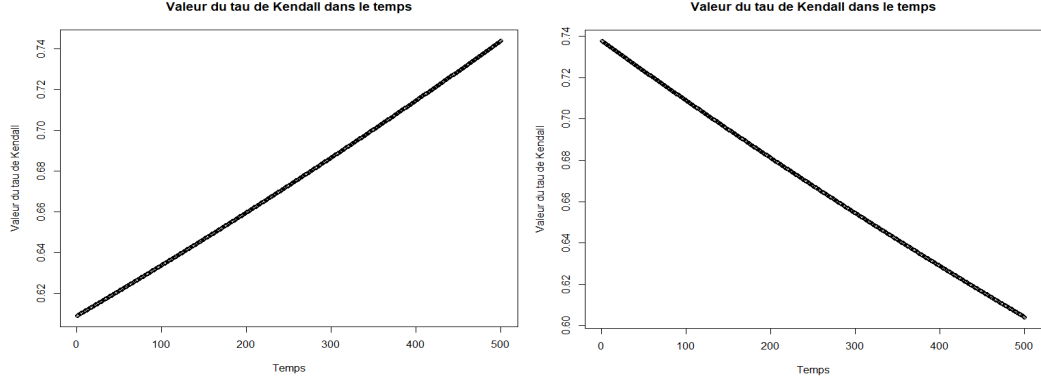


Figure 4.1 – Graphiques de la valeur du tau de Kendall pour $\tau_t = \exp\left(\beta_0 + \beta_1 \cdot \frac{t}{T}\right)$. À gauche, $\beta_0 = -0.5$, $\beta_1 = 0.2$ et $T = 500$. À droite, $\beta_0 = -0.3$, $\beta_1 = -0.2$ et $T = 500$.

définition, il en sera de même pour la valeur de θ_t . La fonction τ_t choisi a donc un impact majeur sur l'estimation des paramètres. On peut observer les valeurs du tau de Kendall à travers le temps pour nos deux derniers tableaux à la Figure 4.1. Le graphique de gauche représente la valeur de la dépendance pour les estimations du tableau 4.4 où β_1 est positif, alors que celui de droite représente la dépendance pour les paramètres du tableau 4.5.

Tout comme pour nos tableaux 4.1, 4.2 et 4.3 représentant l'effet de la variation du paramètre β_1 sur différentes tailles d'échantillon pour nos séries, il peut être intéressant de voir si les résultats seront semblables pour nos tableaux 4.4 et 4.5 si l'on diminue la valeur de T . Les tableaux 4.6 et 4.7 représentent $N = 500$ séries de $T = 150$ observations, pour les mêmes paramètres β_0 et β_1 que nos tableaux précédents.

On remarque très facilement que les estimateurs de nos paramètres sont moins précis. Les valeurs estimées de β_0 sont encore plus petites que la valeur réelle recherchée, mais les valeurs ont un plus grand écart. Les valeurs estimées de β_1 semblent bonnes pour toutes les copules sauf la copule Clayton, mais l'intervalle de confiance est beaucoup plus grand, et inclus maintenant des valeurs positives et négatives à 95% de confiance.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
Gauss	-0.5451	0.2091	[-0.7868,-0.3033]	[-0.0543,0.4725]	[-0.8364,-0.3240]	[-0.0572,0.5072]
	-0.5412	0.2031	[-0.7795,-0.3030]	[-0.0683,0.4745]	[-0.8100,-0.3392]	[-0.0712,0.4769]
Student	-0.5701	0.2149	[-0.8716,-0.2686]	[-0.1194,0.5492]	[-0.9303,-0.3351]	[-0.1086,0.5996]
	-0.5661	0.2108	[-0.8476,-0.2845]	[-0.0971,0.5187]	[-0.8750,-0.3117]	[-0.0810,0.5476]
Gumbel	-0.5917	0.2298	[-0.9226,-0.2609]	[-0.1467,0.6063]	[-0.9699,-0.2988]	[-0.1393,0.6504]
	-0.5317	0.2072	[-0.7205,-0.3428]	[0.0208,0.3936]	[-0.7439,-0.3600]	[0.0163,0.4155]
Clayton	-0.7039	0.2794	[-1.2280,-0.1798]	[-0.3399,0.8987]	[-1.1938,-0.1648]	[-0.3769,0.8414]
	-0.7132	0.2731	[-1.2767,-0.1498]	[-0.3990,0.9453]	[-1.3050,-0.1718]	[-0.4233,0.8620]
Frank	-0.5393	0.1827	[-0.7839,-0.2947]	[-0.1505,0.5159]	[-0.8132,-0.3378]	[-0.1282,0.4927]
	-0.5380	0.1911	[-0.7708,-0.3052]	[-0.1284,0.5105]	[-0.8197,-0.3343]	[-0.1272,0.5227]

Tableau 4.6 – Estimation pour différentes familles de copule obtenue à l’aide de $N = 500$ séries de $T = 150$ observations, pour des paramètres $\beta_0 = -0.5$ et $\beta_1 = 0.2$

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	I.C. (β_0) (N)	I.C. (β_1) (N)	I.C. (β_0) (Q)	I.C. (β_1) (Q)
Gauss	-0.3740	-0.2010	[-0.5584,-0.1895]	[-0.5060,0.1040]	[-0.5917,-0.2222]	[-0.5319,0.1183]
	-0.3693	-0.2016	[-0.5421,-0.1965]	[-0.5017,0.0985]	[-0.5672,-0.2280]	[-0.5210,0.0809]
Student	-0.3801	-0.2082	[-0.5867,-0.1735]	[-0.5699,0.1535]	[-0.6171,-0.2104]	[-0.5651,0.1511]
	-0.3758	-0.2132	[-0.5737,-0.1778]	[-0.5552,0.1288]	[-0.6073,-0.2221]	[-0.6032,0.0980]
Gumbel	-0.3990	-0.2247	[-0.6382,-0.1598]	[-0.6530,0.2037]	[-0.6675,-0.1970]	[-0.6871,0.1694]
	-0.3869	-0.2265	[-0.6167,-0.1572]	[-0.6417,0.1887]	[-0.6472,-0.1969]	[-0.6738,0.1948]
Clayton	-0.4589	-0.2555	[-0.8065,-0.1112]	[-0.8683,0.3573]	[-0.8539,-0.1610]	[-0.8502,0.4130]
	-0.4633	-0.2410	[-0.8393,-0.0874]	[-0.8357,0.3538]	[-0.8904,-0.1553]	[-0.8493,0.3283]
Frank	-0.3729	-0.1943	[-0.5507,-0.1950]	[-0.5011,0.1125]	[-0.5824,-0.2217]	[-0.5137,0.0890]
	-0.3704	-0.1935	[-0.5537,-0.1871]	[-0.5178,0.1309]	[-0.5892,-0.2260]	[-0.4959,0.1411]

Tableau 4.7 – Estimation pour différentes familles de copule obtenue à l’aide de $N = 500$ séries de $T = 150$ observations, pour des paramètres $\beta_0 = -0.3$ et $\beta_1 = -0.2$

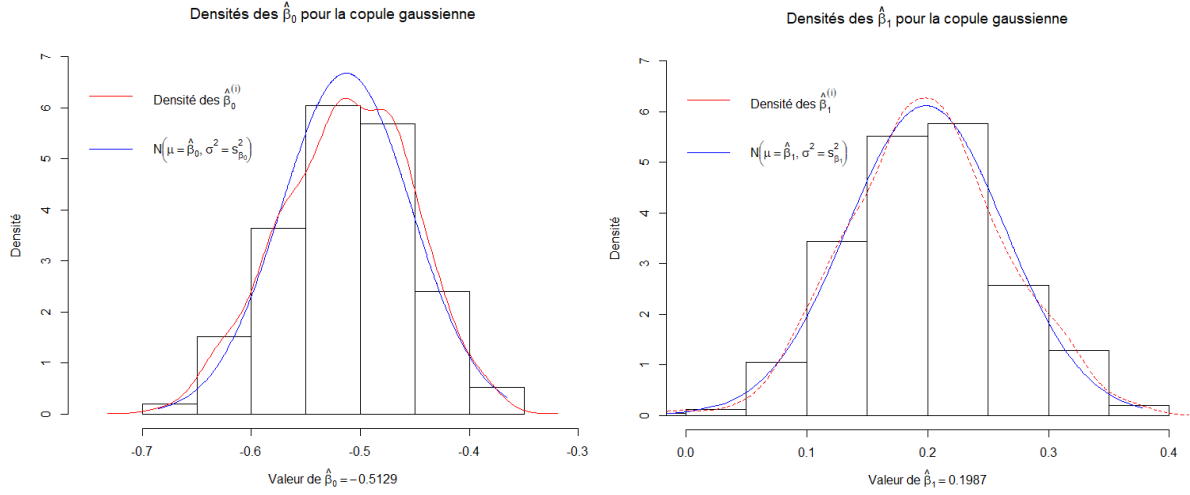


Figure 4.2 – Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule gaussienne du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$.

Il est donc avisable d'affirmer que la dépendance change à travers le temps pour ces paramètres, lorsque $T = 150$.

La valeur de l'estimation pour β_1 ainsi qu'un intervalle de confiance pour sa validité nous donne déjà un bon point de vue initial sur notre méthode d'estimation. Par contre, rappelons que $\hat{\beta}_1$ est calculé comme une moyenne des $\hat{\beta}_1^{(i)}$ de chaque série, c'est-à-dire pour $i = 1, \dots, N$. Il serait intéressant de regarder si les N valeurs obtenues semblent suivre une loi normale. Notons d'ailleurs qu'un des deux intervalles de confiance utilisé est celui de la loi normale et qu'il ne serait pas valide si les données testées ne correspondaient pas à ce dernier. Pour ce faire, cinq estimations furent sélectionnées du tableau 4.4, soit la première des deux lignes pour chacune des copules. On obtient le graphique des densités pour chacune de ces copules aux Figures 4.2 à 4.6.

Pour chacun des graphiques aux figures précédentes, la courbe en rouge représente la densité des $N = 500$ valeurs estimées de nos paramètres β_0 et β_1 . On peut comparer cette courbe avec celle en bleu représentant la densité d'une loi normale ayant pour

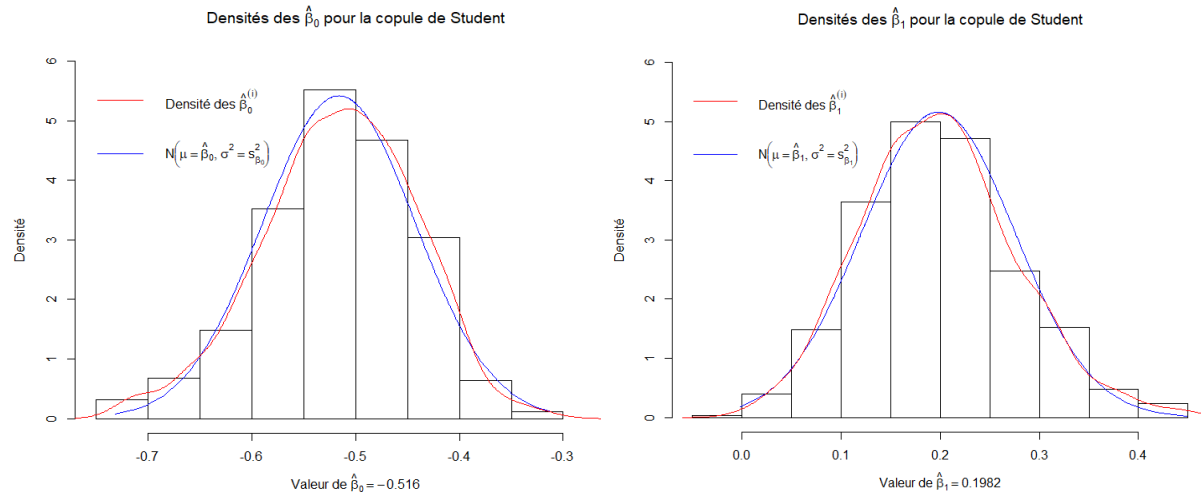


Figure 4.3 – Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Student du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$.

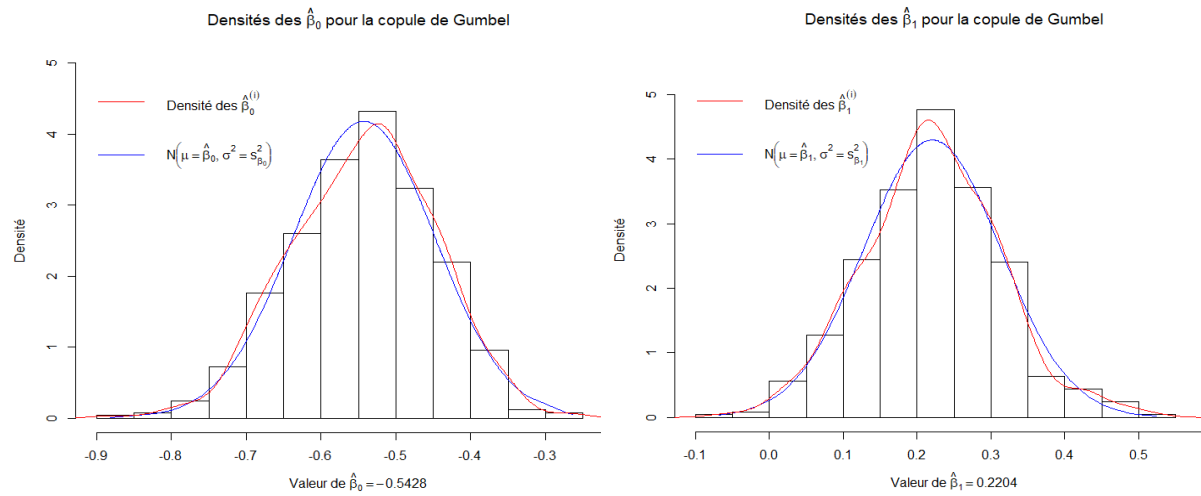


Figure 4.4 – Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Gumbel du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$.

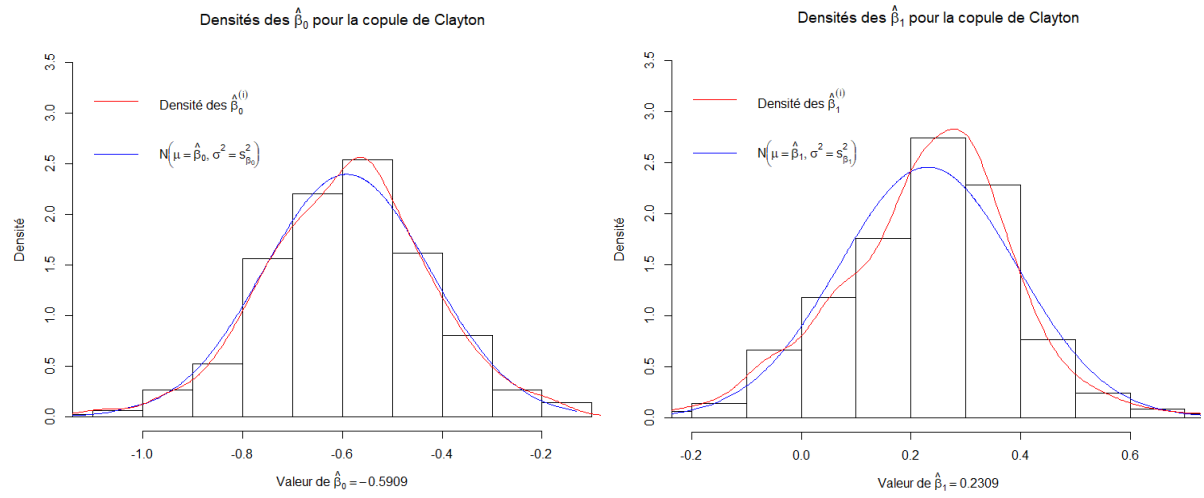


Figure 4.5 – Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Clayton du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$.

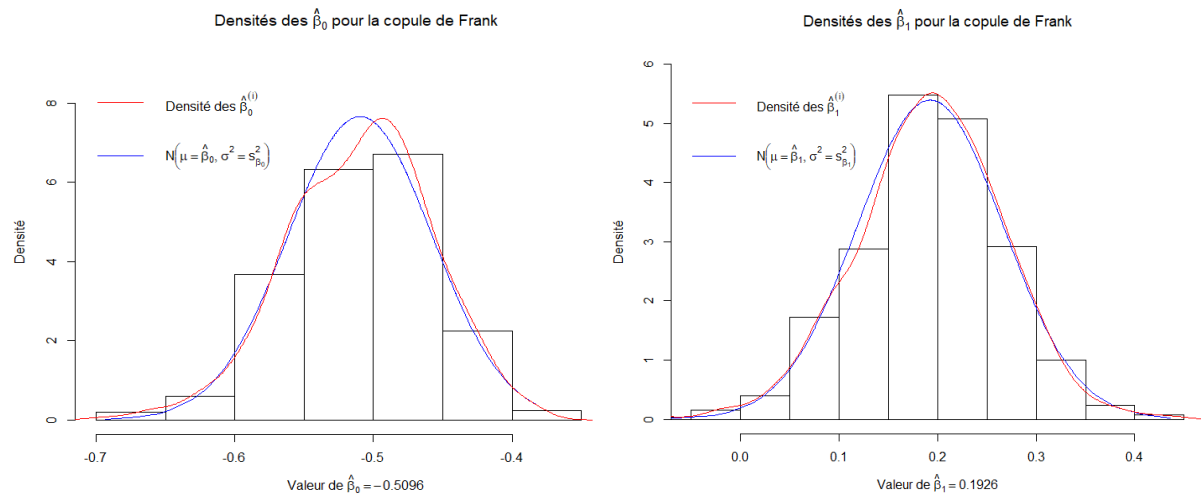


Figure 4.6 – Graphiques de la densité des paramètres $\hat{\beta}_0$ et $\hat{\beta}_1$ pour la première des deux simulations de la copule de Frank du tableau 4.4. $T = 500$ observations, $\beta_0 = -0.5$ et $\beta_1 = 0.2$.

paramètre μ la moyenne des $N = 500$ valeurs estimées de β_0 (resp β_1) et σ l'écart-type de ces 500 valeurs.

Les tests de normalité de Kolmogorov-Smirnov et de Cramer-Von Mises indiquent que les distributions empiriques des estimateurs obtenus aux Figures 4.2 à 4.6 concordent avec des lois normales, ce qui est un résultat attendu étant donné la convergence asymptotique démontrée à la section 3.3.

4.2 Application empirique

Le jeu de données analysé provient du site officiel de la FAO (*Food and Agriculture Organization*). On y retrouve les prix mensuels du blé entre le mois de janvier 2000 et juillet 2017 pour des pays tels que l'Australie, le Royaume-Uni, le Kazakhstan, le Canada ainsi que la région de la Mer Noire entre l'Europe, le Caucase et l'Anatolie. Sur cette période, certaines données sont manquantes, ce qui pose problème pour la construction des séries chronologiques. Ainsi, pour chacun des cinq pays sélectionnés, la série chronologique analysée couvre la période la plus longue sans coupure dans les données. Cela a pour conséquences que les échantillons sélectionnés ne possèdent pas tous la même taille, et ne couvre pas tous la même période.

Les séries de prix du blé (mois initial, mois final, taille de l'échantillon) sont définies comme suit : l'Australie (décembre 2008, juillet 2017, 104) ; le Royaume-Uni (juillet 2008, juillet 2017, 109) ; le Kazakhstan (octobre 2006, juillet 2017, 130) ; le Canada (août 2000, avril 2012, 141) ; et la Mer Noire (juillet 2004, octobre 2010, 76).

L'estimation des paramètres se fera de la même façon qu'à la section précédente.

Toutefois, l'**Algorithme 1** ne sera pas nécessaire étant donné que les séries chronologiques existent, et donc seule l'étape 1 de l'**Algorithme 2** sera utilisée.

Comme des valeurs différentes des paramètres β_0 et β_1 sont obtenues pour les familles de copules choisies, les critères d'information AIC et BIC seront appliqués afin de déterminer la famille de copule la plus appropriée pour chacune des cinq séries de prix du blé. La famille de copule correspondant à la valeur minimale des critères sera sélectionnée. Les résultats sont présentés aux tableaux 4.8 à 4.12.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	AIC	BIC
Gaussienne	-0.2475	0.0078	-215.667	-210.378
Student	-0.2193	-0,0411	-219.655	-214.366
Gumbel	-0.2966	0.1002	-219.254	-213.965
Clayton	-0.2809	-0.2152	-162.623	-157.334
Frank	-0.2528	0.0682	-220.544	-215.255

Tableau 4.8 – Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé de l'Australie, $T = 104$.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	AIC	BIC
Gaussienne	-0.4242	0.3311	-233.141	-227.758
Student	-0.4659	0.3727	-226.479	-221.097
Gumbel	-0.4953	0.4236	-231.568	-226.185
Clayton	-0.7551	0.5810	-162.343	-156.100
Frank	-0.3356	0.2297	-246.627	-241.244

Tableau 4.9 – Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé du Royaume Uni, $T = 109$.

La copule la plus appropriée pour les séries de prix du blé de notre application empirique est la copule de Frank trois fois, la copule gaussienne une fois et la copule de Student une fois. Il est important de faire un lien entre la longueur des séries et les résultats des applications théoriques de la Section 4.1. Comme l'écart-type des estimés des paramètres augmente lorsque la taille de l'échantillon diminue, et nos séries comptent entre 76 et 141 observations, la précision de nos résultats ici rejoint celle des tableaux 4.6 et 4.7, où $T = 150$. Notons que la pour la série du prix du blé du Canada,

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	AIC	BIC
Gaussienne	-0.3729	0.2732	-295.155	-289.420
Student	-0.3665	0.2822	-305.961	-300.226
Gumbel	-0.4617	0.3960	-288.267	-282.532
Clayton	-0.4645	0.3117	-257.977	-252.242
Frank	-0.3033	0.1828	-293.538	-287.803

Tableau 4.10 – Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé du Kazakhstan, $T = 130$.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	AIC	BIC
Gaussienne	-0.2223	0.1265	-407.369	-401.471
Student	-0.1890	0.0829	-414.999	-409.101
Gumbel	-0.1255	-0.0687	-392.704	-386.807
Clayton	-0.2454	0.2176	-389.602	-383.705
Frank	-0.1778	0.1074	-442.770	-436.873

Tableau 4.11 – Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé du Canada, $T = 141$.

Copule	$\hat{\beta}_0$	$\hat{\beta}_1$	AIC	BIC
Gaussienne	-0.1785	-0.0435	-179.469	-174.808
Student	-0.2100	0,0000	-177.625	-172.963
Gumbel	-0.1149	-0,1972	-178.100	-173.438
Clayton	-0.4503	0.2744	-143.897	-139.236
Frank	-0.1137	-0.1408	-178.901	-174.239

Tableau 4.12 – Valeurs estimées des paramètres des cinq copules pour leur famille respective, pour la série du prix du blé de la région de la Mer Noire, $T = 76$.

la copule de Clayton n'offrait aucun résultat avec notre méthode d'estimation. Pour résoudre ce problème, la simulation de cette copule a été fait en deux étapes. L'estimation du paramètre β_0 a d'abord été effectuée, puis à partir de $\hat{\beta}_0$, le paramètre β_1 a ensuite été estimé.

Les valeurs de $\hat{\beta}_1$ pour le Royaume Uni et le Kazakhstan montrent une différence marquée par rapport au scénario où la dépendance ne serait pas fonction du temps ($\beta_1 = 0$). Les valeurs obtenues de $\hat{\beta}_1$ pour l'Australie, le Canada ainsi que la région de la Mer Noire sont différentes de 0, sans toutefois être trop loin. De plus, les valeurs de β_1 estimées pour le Royaume-Uni et le Kazakhstan sont positives, suggérant que la dépendance dans les deux séries augmente avec le temps. On peut l'observer en regardant la valeur du tau de Kendall à travers le temps pour ces deux pays à la Figure 4.7.

Dans les deux cas, le tau de Kendall débute entre 0.69 et 0.72 et augmente pour atteindre des valeurs entre 0.89 et 0.92.

Copule	$\hat{\beta}_0$	AIC	BIC
Gaussienne	-0.2435	-217.649	-215.004
Student	-0.2402	-221.317	-218.672
Gumbel	-0.2451	-219.697	-217.053
Clayton	-0.3970	-157.976	-155.332
Frank	-0.2185	-221.334	-218.690

Tableau 4.13 – Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé de l'Australie, $T = 104$.

Afin de comparer les résultats d'une copule à paramètre dynamique versus un paramètre statique, des estimations ont été obtenues pour l'équation 3.4 où β_1 est posé à 0. L'équation devient $\tau_t^* = \exp(\beta_0)$. Pour les cinq séries du prix du blé, une estimation $\hat{\beta}_0$ a été obtenue pour τ_t^* . On retrouve ces résultats aux tableaux 4.13 à 4.17.

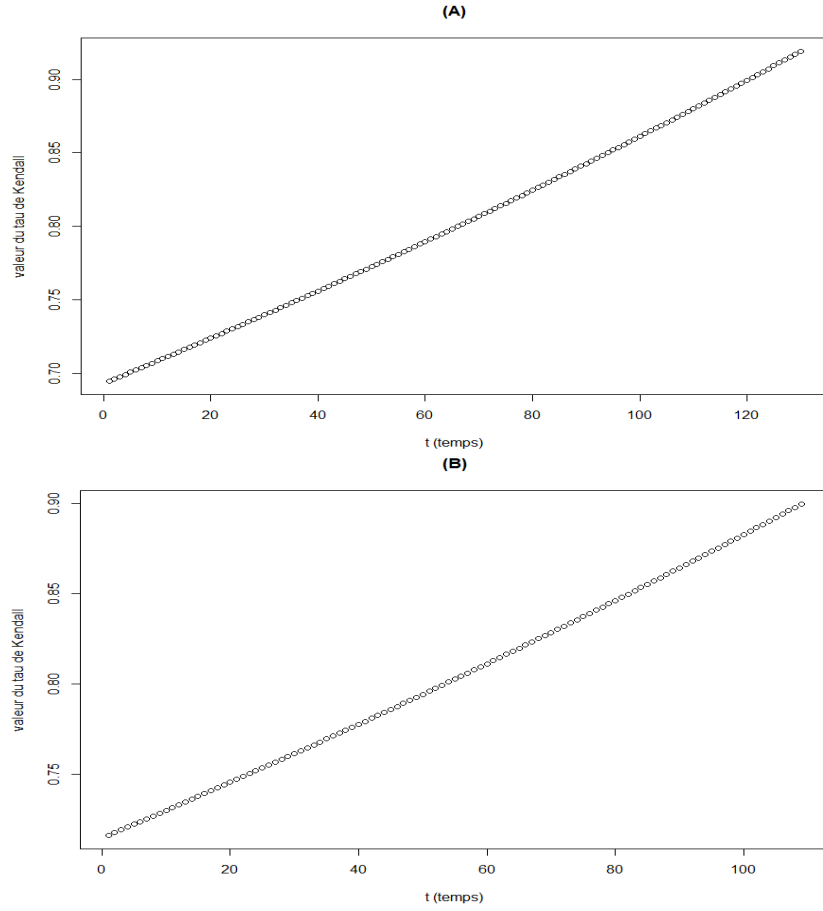


Figure 4.7 – Valeurs du tau de Kendall pour $\tau_t = \exp\left(\beta_0 + \beta_1 \cdot \frac{t}{T}\right)$, avec β_0 et β_1 les valeurs estimées pour (A), la copule de Student (Kazakhstan) et (B), la copule de Frank (Royaume-Uni).

Copule	$\hat{\beta}_0$	AIC	BIC
Gaussienne	-0.2831	-201.522	-198.831
Student	-0.2754	-205.158	-202.467
Gumbel	-0.2819	-206.582	-203.890
Clayton	-0.4940	-128.987	-126.295
Frank	-0.2152	-232.561	-229.869

Tableau 4.14 – Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé du Royaume-Uni, $T = 109$.

Copule	$\hat{\beta}_0$	AIC	BIC
Gaussienne	-0.2411	-275.742	-272.875
Student	-0.2181	-291.452	-288.584
Gumbel	-0.2518	-264.876	-262.008
Clayton	-0.3102	-244.368	-241.500
Frank	-0.2061	-286.921	-284.053

Tableau 4.15 – Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé du Kazakhstan, $T = 130$.

Copule	$\hat{\beta}_0$	AIC	BIC
Gaussienne	-0.1617	-396.201	-393.253
Student	-0.1455	-412.541	-409.593
Gumbel	-0.1618	-392.264	-389.315
Clayton	-0.2455	-313.120	-310.171
Frank	-0.1211	-435.655	-432.706

Tableau 4.16 – Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé du Canada, $T = 141$.

Copule	$\hat{\beta}_0$	AIC	BIC
Gaussienne	-0.2010	-180.933	-178.602
Student	-0.2100	-179.625	-177.294
Gumbel	-0.2161	-175.122	-172.791
Clayton	-0.3281	-136.656	-134.325
Frank	-0.1898	-176.115	-173.784

Tableau 4.17 – Valeurs du paramètre de la copule pour leur famille respective, pour les prix du blé de la Mer Noire, $T = 76$.

On remarque que les critères du AIC et du BIC propose les mêmes familles de copules pour la cas statique que pour le cas où la dépendance changeait à travers le temps. De plus, on obtient des valeurs inférieures pour l'AIC et le BIC quand le paramètre est constant pour l'Australie et la Mer Noire. Dans le cas dynamique, les valeurs de β_1 estimées pour ces deux endroits étaient très près de 0. Ces résultats ne sont donc pas inattendus.

CONCLUSION

La structure de dépendance d'une série chronologique est depuis longtemps un sujet d'étude très populaire, et cela est d'autant plus vrai depuis l'apparition des copules. La copule est un des concepts les plus utilisés dans le domaine de l'analyse de la dépendance. Elle nous offre une réponse beaucoup plus informative qu'un simple chiffre fourni par un coefficient de corrélation ou d'autocorrélation. La modification du paramètre de la copule statique vers un paramètre variant dans le temps permet à la copule d'obtenir une flexibilité supplémentaire.

Dans ce mémoire, la structure de dépendance d'une série chronologique a été étudiée à l'aide des copules. Le cas classique des copules à paramètre statique a servi de point de départ, ce qui représente bien une série chronologique stationnaire. Or, la force de dépendance dans les séries chronologiques peut être croissante ou décroissante dans le temps. Afin de modéliser ces séries non-stationnaires, une fonction représentant le paramètre monotone (soit croissant, soit décroissant) d'une copule dynamique a été proposé pour chacune des familles. La principale qualité de cette fonction est d'être en lien avec le tau de Kendall. Ainsi, la comparaison des paramètres des différentes familles de copules fut possible. De plus, la fonction choisie respecte, pour chacune des copules, le domaine respectif du paramètre. Des méthodes d'estimation ainsi que des algorithmes de simulation ont été étudiées. Aussi, nous avons établi la convergence des

estimateurs. Une application à des séries du prix du blé a permis de concrétiser cette avancée méthodologique, en plus d'une application théorique.

Une fonction permettant à la dépendance d'une série chronologique d'augmenter et de diminuer dans le temps permettrait une plus grande flexibilité au niveau du modèle.

L'estimation de paramètres convergents d'une telle fonction représente un défi intéressant. D'ailleurs, des tests pourraient être un ajout intéressant dans le but de tester si la dépendance est significativement dépendante ou stable dans le temps. Enfin, l'étude de la dépendance entre deux séries chronologiques possédant une structure de dépendance dynamique est une autre piste à suivre.

Bibliographie

- [Aka74] Hirotugu Akaike. A New Look at the Statistical Model Identification. *IEEE transactions on automatic control*, 19(6) : 716–723, 1974.
- [BD91] P.J. Brockwell and R.A. Davis. *Time Series : Theory and Methods*. Springer-Verlag, New York, 1991.
- [BD02] P.J. Brockwell and R.A. Davis. *Introduction to Times Series and Forecasting*. Springer, Berlin, 2002.
- [BJ76] G.E.P. Box and G.M. Jenkins. *Time Series Analysis : Forecasting and Control*. Holden-Day series in time series analysis and digital processing. Holden-Day, 1976.
- [BJ90] George Edward Pelham Box and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated, 1990.
- [Bol86] Tim Bollerslev. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3) : 307–327, 1986.
- [CF06] Xiaohong Chen and Yanqin Fan. Estimation and Model Selection of Semiparametric Copula-based Multivariate Dynamic Models under Copula Misspecification. *Journal of Econometrics*, 135(1–2) : 125–154, 2006.
- [CH05] Song Xi Chen and Tzeeming Huang. Nonparametric Estimation of Copula Functions for Dependence Modeling. 2005.

- [Che99] S.X. Chen. Beta Kernel Estimators for Density Functions. *Computational Statistics and Data Analysis*, 31 : 131–145, 1999.
- [Che00] S.X. Chen. Probability Density Function Estimation using Gamma Kernels. *Ann. Inst. Statist. Maths*, 52 : 471–480, 2000.
- [CL98] G. Casella and E.L. Lehmann. *Theory of Point Estimation*. Springer, New York, 1998.
- [Deh80] P. Deheuvels. Non Parametric Tests for Independence. *Springer, Lecture Notes in Mathematics*, 821, 1980.
- [Eng82] Robert Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4) : 987–1007, 1982.
- [FFK02] Hong-Bin Fang, Kai-Tai Fang, and Samuel Kotz. The Meta-elliptical Distributions with Given Marginals. *Journal of Multivariate Analysis*, 82 : 1–16, 2002.
- [FRW02] J.D Fermanian, D. Radulovic, and M. Wegkamp. Weak Convergence of Empirical Copula Processes. (2002-06), 2002.
- [FV98] Edward W. Frees and Emiliano A. Valdez. Understanding Relationships using Copulas. *N. Am. Actuar J*, 2(1) : 1–25, 1998.
- [GGR95] C. Genest, K. Ghoudi, and L.-P. Rivest. A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions. *Biometrika*, 82(3) : 543–552, 1995.
- [GM90] I. Gijbels and J. Mielniczuk. Estimation of the Density of a Copula Function. *Communications in Statistics*, 19 : 445–464, 1990.
- [GOV09] I. Gijbels, M. Omelka, and N. Veraverbeke. Improved Kernel Estimation of Copulas : Weak Convergence and Goodness-of-fit Testing. *Annals of Statistics*, 37, 2009.

- [GR04] C. Genest and B. Remillard. Test of Independence and Randomness Based on the Empirical Copula Process. *Test*, 13 : 335–369, 2004.
- [Gra69] C.W.J. Granger. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3) : 424–438, 1969.
- [Gre14] Ievgenii Grebennikov. Moment-based Estimation of Dependence Parameters in Copula Models for Discrete Data. 2014.
- [Hos81] J.R.M. Hosking. Fractional Differencing. *Biometrika*, 68 : 165–176, 1981.
- [Joe97] Harry Joe. *Multivariate Models and Dependence Concepts*. volume 73 of Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1997.
- [Joe05] Harry Joe. Asymptotic Efficiency of the Two-Stage Estimation Method for Copula-based models. *J. Multivariate Anal.*, 94(2) :401–419, 2005.
- [JX96] H. Joe and J. Xu. The Estimation Method for Inference Functions for Margins for Multivariate Models. *Technical Report*, 166, 1996.
- [KSS07] Gunky Kim, Mervyn J. Silvapulle, and Paramsothy Silvapulle. Comparison of Semiparametric and Parametric Methods for Estimating Copulas. *Computational Statistics and Data Analysis*, 51(6) : 2836–2850, 2007.
- [Nel99] R.B. Nelsen. *An Introduction to Copulas*. Lecture notes in statistics. Springer, 1999.
- [NRB19] B.R. Nasri, B.N. Remillard, and T. Bouezmarni. Semi-parametric Copula-based Models under Non-stationarity. *Journal of Multivariate Analysis*, 173 : 347–365, 2019.
- [Pat09] A.J. Patton. *Copula-Based Models for Financial Time Series*. Handbook of Financial Time Series. Springer, Berlin, 2009.
- [Pat12] Andrew J. Patton. A Review of Copula Models for Economic Time Series. *Journal of Multivariate Analysis*, 110 : 4–18, 2012.

- [RPS12] Bruno Remillard, Nicolas Papageorgiou, and Frederic Soustra. Copula-based Semiparametric Models for Multivariate Time Series. *Journal of Multivariate Analysis*, 110 : 30–42, 2012.
- [Sch78] Gideon Schwarz. Estimating the Dimension of a Model. *Ann. Statist.*, 6(2) : 461–464, 03 1978.
- [Sim80] C.A. Sims. Macroeconomics and Reality. *Econometrica*, 48(1) : 1–48, 1980.
- [Sk159] Abe Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8 :229–231, 1959.
- [SW81] B. Schweizer and E.F. Wolff. On Nonparametric Measures of Dependence for Random Variables. *Annals of Statistics*, 9(4), 1981.
- [Wal31] Sir Gilbert Walker. On Periodicity in Series of Related Terms. *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, 131(818) :518–532, 1931.
- [Whi51] Peter Whittle. *Hypothesis Testing in Time Series Analysis*. 1951.
- [Yul27] G. Udny Yule. On a Method of Investigating Periodicities in Disturbed Series, with Special reference to Wolfer’s Sunspot Numbers. *Philosophical Transactions of the Royal Society of London Series A*, 226 :267–298, 1927.